Master in Actuarial and Financial Sciences

MASTER THESIS

Title: Credit risk measures and the estimation error in the ASRF model under the Basel II IRB approach

Author: Aiberson Ventura Horán

Advisor: Luis Ortiz Gracia

Academic year: 2023-2024

UNIVERSITAT DE BARCELONA
Facultat d'Economia i Empresa

Màster de Ciències Actuarials i Financeres

Faculty of Economics and Business

Universitat de Barcelona

Master thesis

Master in Actuarial and Financial Sciences

CREDIT RISK MEASURES AND THE ESTIMATION ERROR IN THE ASRF MODEL UNDER THE BASEL II IRB APPROACH

Author: Aiberson Ventura Horán

Advisor: Luis Ortiz Gracia

The content of this document is the sole responsibility of the author, who declares that he has not incurred plagiarism and that all references to other authors have been expressed in the text.

Abstract

This project aims to investigate the impact of estimating the probability of default, due to its unknown true value, to estimate the Value-al-Risk under the ASRF model. To accomplish this, a Monte Carlo simulation approach is employed to estimate de default ratio based on predetermined ”real” probabilities of default. By simulating different scenarios, we can assess the potential bias and evaluate the need for adjustments, such as confidence level modifications or the inclusion of a Margin of Conservatism (MoC), to account for estimation uncertainty.
Keywords: *Estimation error, Probability of Default, ASRF model,Value at Risk, Monte Carlo simulation*.

Contents

1 Introduction **4**
1.1 Real value vs estimation 4
1.2 The BCBS decision 5

2 Theoretical foundation of the model **7**
2.1 Creditworthiness under ASRF Model 7
2.2 Probability of default 8
2.3 From theory to practice 9
2.3.1 Introducing the binomial distribution 9
2.3.2 Conditional default rate 11
2.3.3 Unconditional default rate 12
2.3.4 Using DR_t to estimate PD^{LR} 15
2.3.5 A solution to correct the underestimation error 16

3 Monte Carlo's theoretical foundation **18**
3.1 Random variables and the law of large numbers 18
3.2 Variance reduction 19

4 Application and results **24**
4.1 Implementation in R 24
4.2 Results 25

5 Conclusions **28**

A R code **29**

Chapter 1

Introduction

The computation of risk measures in the field of financial modeling is subject to various sources of uncertainty, including parameter estimation errors, which is currently not accounted for. One critical parameter in risk assessment is the Long Run Probability of Default (PD^{LR}) within the ASRF (Asymptotic Single Risk Factor) model and its true value is unknown. Thus, it is necessary to use the observed default ratio to estimate it. This project is motivated by the need to understand the impact of the error of estimating the probability of default in the computation of the widely used Value at Risk (VaR). In order to do so, we will simulate different scenarios based on predetermined real values and compare the quantiles computed with the real PD^{LR} and with the estimated one. We will see that we are actually underestimating the risk by using an estimation, and therefore a possible solution will be evaluated.

1.1 Real value vs estimation

Over fifteen years ago, the Basel Committee on Banking Supervision (BCBS)[1] implemented the Basel II risk framework, by which banks were allowed to use their own internal models to calculate minimum capital requirements for different types of risks. However, they mandated a specific model for credit risk, so banks could only estimate certain parameters of the model. The goal of this model is to provide a vehicle to estimate the maximum loss, in case the bank can not meet its own credit obligations by resorting to its capital. Given a certain confidence level (α), risk measures such as Value-at-Risk (VaR_α)[2] and Expected Shortfall (ES_α)[3] can be computed using the Merton-Vasicek model, as it can be seen in Huang and Oosterlee (2007) and Osmundsen (2018). However, in order to do so, the estimation of a parameter is required since the true value is unknown. That is to say, that we need to replace the ”true value” of the parameter, which appears in the theoretical model, by some estimation obtained from observed data.

Estimations always introduce variability and potential error since they are derived from a

[1] See Basel Committee (2005).

[2] The VaR_α is just the quantile whose probability of seeing a greater value is α.

[3] ES_α is the average loss, given that the loss is greater than the VaR_α.

sample of data, real or simulated, and the specific observations in the sample may not perfectly represent the entire population, leading to sampling variability. This means that different samples could yield different estimations, introducing uncertainty into the parameter substitution.[4]
Therefore, unless this new source of uncertainty is taken into account, we would be assuming a greater error than what we expect. As said before, standard risk measures such as VaR_α have a confidence level α which establishes the error we are willing to accept.
By definition, the error we are willing to accept is $1-\alpha$. However, let us say that the loss is a function that depends on some parameter θ, so the Value-at-Risk also depends on this parameter ($\text{VaR}^\theta_\alpha(L)$). If the real value of the parameter is known, we can be sure that we are actually accepting a $1-\alpha$ error, but if the parameter is unknown and, thus, must be estimated, the probability of surpassing the Value-at-Risk might be higher than $1-\alpha$ owing to the inherent variability of the estimation of θ.

1.2 The BCBS decision

The Basel Committee on Banking Supervision (BCBS) utilizes VaR_α measures to determine the amount of equity capital required. Since estimating the parameters of the model induces another source of error[5], the BCBS sets a higher confidence level α to account for this additional error. As an illustration, let us consider the target confidence level $\alpha = 99\%$ which means that the loss can't be larger than the Value-at-Risk more than 1% of the times. As said in the previous section, since the parameter needs to be estimated, it is likely that the Value-at-Risk ($\widehat{\text{VaR}}_{99\%}$) obtained with the estimator $\hat{\theta}$ would not ensure that the realized loss exceeds this threshold with a probability 1%. That is to say that $\widehat{\text{VaR}}_{99\%}$) underestimates the $\text{VaR}_{99\%}$ computed with the real parameter.
The BCBS, acknowledging this, tries to correct this bias by increasing the confidence level α to, for instance, 99.9% . Nevertheless, in compliance with European regulations, banks utilizing internal models are obligated to integrate an additional Margin of Conservatism (MoC) that should be proportional to the estimation error. Moreover, as per the European Banking Authority (EBA), the quantification of the MoC for general estimation errors should mirror the dispersion of the statistical estimator's distribution as it can be seen on EBA (2017). From this two possible options can be inferred:

1. The rectification proposed by the BCBS, such as establishing the confidence level at 99.9%, is regarded as inadequate in accordance with European regulations.

2. The desired confidence level in Europe is indeed 99.9% rather than 99%.

This study aims to evaluate whether a confidence level of 99.9% is sufficient to offset the estimation error that would be mitigated with a confidence level of 99% or 99.5%, assuming

[4]In general, there also other sources of error such as model assumptions that are not met by the data or measurement errors.

[5]See Casellina et al. (2021)

the first interpretation. Furthermore, a potential methodology for introducing a MoC to tackle the estimation error will also be explored.

The project's structure is delineated as follows: In Chapter 1, we introduce the problem at hand an the motivation behind it, along with a short overview of the project's layout. Chapter 2 will delve into the theoretical underpinnings of the ASRF model and the computation of PD^{LR} essential for VaR determination, while also introducing the solution proposed by Casellina et al. (2023). Chapter 3 will furnish the required background for employing Monte Carlo methods in this context, along with a comprehensive elucidation of two prominent variance reduction techniques. The first one is importance of sampling, which is used in Casellina et al. (2023), but not explained, so we had to investigate the complex mathematical background behind it and figure out how to apply it to our problem. The second one is the technique of Antithetic variate, which is our contribution to Casellina's work. This is a less complex variance reduction technique, but improved the speed at which we get to the real value which is the main problem we have when using Monte Carlo. Subsequently, Chapter 4 will apply the aforementioned theoretical framework and discuss the outcomes. Finally, in Chapter 5, conclusions will be drawn, offering perspectives on both the Margin of Conservatism proposed by regulators and Casellina's solution.

Chapter 2

Theoretical foundation of the model

2.1 Creditworthiness under ASRF Model

The standard theoretical framework adopted by the BCBS is the Asymptotic Single Risk Factor (ASRF) model. In this paper, we disregard some aspects of the Supervisory Formula[1] such as the maturity adjustment and we assume the asset correlation between borrowers' assets ω is a fixed parameter[2], while in the BCBS framework it varies as a function of probability of default.

In the ASRF framework, following the Merton-Vasicek model[3], the i-th creditworthiness change is defined as a function of two random variables:

$$Y_{i,t} = \sqrt{\omega} \cdot Z_t + \sqrt{1-\omega} \cdot W_{i,t} \quad \forall i = 1, 2, ...N \tag{2.1}$$

N is the portfolio size. The two random variables are Z_t and $W_{i,t}$. The first one is the common factor, which affects every borrower and shows the state of the economy, and the later is the idiosyncratic factor, which is different for each borrower. It is also worth noticing that $\omega \in (0, 1)$, which establishes the correlation between Z_t and $Y_{i,t}$, is an exogenous parameter set by regulation.

The hypotheses are the following:

$HP.1 : W_{i,t} \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$.
$HP.2 : Z_t \sim \mathcal{N}(0, 1)$.
$HP.3 : Corr(W_{i,t}, Z_t) = 0$.
$HP.4 : Corr(Z_{t-1}, Z_t) = 0$.
$HP.5 : Y_{i,t}$ is linear.
$HP.6$: The portfolio is infinitely grained.

The first three Hypotheses show that the two random variables affecting the creditworthi-

[1]This is the approach proposed by the Basel Committee to calculate regulatory capital.
[2]For some asset classes, like residential mortgages, the asset correlation is constant.
[3]Based on the widely used Gaussian copula model. See Glasserman (2005).

ness are standard normal distribution and uncorrelated. HP.4 is not usually mentioned in the standard presentation of the ASRF model because it is not necessary so long we are only dealing with one period. However, since we will later introduce an estimator of a parameter of the model as the average of $t = 1, 2...T$ observations, this is included for the sake of simplicity, since this hypothesis excludes any serial correlation. Notice that Z_t and Z_{t-1} are normally distributed and come from the same normal distribution, thus, this hypothesis implies that they are independent. HP.6 means that the number of borrowers is large enough, so the effect on the idiosyncratic factor is negligible[4]. Therefore, Z_t will be the most important factor.

Once we have the i-th creditworthiness change clearly defined, we can infer its probability distribution. Notice that $Y_{i,t}$ is a linear combination of two independent normal distributions, and therefore it also follows a normal distribution. Let's compute its parameters:

$$\mathbb{E}[Y_{i,t}] = \mathbb{E}[\sqrt{\omega} \cdot Z_t + \sqrt{1-\omega} \cdot W_{i,t}] = \sqrt{\omega} \cdot \mathbb{E}[Z_t] + \sqrt{1-\omega} \cdot \mathbb{E}[W_{i,t}] = 0.$$

$$\mathbb{V}[Y_{i,t}] = \mathbb{V}[\sqrt{\omega} \cdot Z_t + \sqrt{1-\omega} \cdot W_{i,t}] = \omega \cdot \mathbb{V}[Z_t] + (1-\omega) \cdot \mathbb{V}[W_{i,t}] = \omega + 1 - \omega = 1.$$

Thus, $Y_{i,t} \sim \mathcal{N}(0, 1)$.

2.2 Probability of default

In this section, we are going to define the probability of default at a specific period t[5] ($PD_{i,t}$), where . That is, the probability that during the time period t the value of the assets is under a threshold s. Formally,

Definition 2.2.1 (Probability of Default at time t)**.** Given a threshold $s \in \mathbb{R}$, the Probability of Default at time t of the i-th borrower($PD_{i,t}$) is defined as:

$$PD_{i,t} := \mathbb{P}[Y_{i,t} < s] = \Phi(s) \tag{2.2}$$

Where $\Phi(s)$ is the cumulative distribution function of the standard normal distribution evaluated at s:

$$\Phi(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{s} e^{\frac{-x^2}{2}} \, dx$$

Notice that $PD_{i,t} = \Phi(s)$ comes from the fact that $Y_{i,t}$ follows an standard normal distribution. $PD_{i,t}$ depends on the time period, therefore it changes at each period. Thus, we define the *long run probability of default* PD^{LR}[6] as an "equilibrium" or steady value for the probability of default. As a consequence,

$$s = \Phi^{-1}(PD^{LR}) \tag{2.3}$$

[4]This will be shown later in this chapter.

[5]Notice that t is not a specific point in time, but an interval $(t-1, t)$

[6]Formally, $PD^{LR} = \mathbb{E}[PD_{i,t}]$.

Coming back to the definition of $PD_{i,t}$, it clearly depends on Z_t, thus let us compute the probability of default at a time period t conditioned to $Z_t = z$. Since, $PD_{i,t} = \mathbb{P}[Y_{i,t} < s]$ the first step should be to determine the distribution of $Y_{i,t}|Z_t = z$:

$$Y_{i,t}^z = (Y_{i,t}|Z_t = z) = \sqrt{\omega} \cdot z + \sqrt{1-\omega} \cdot W_{i,t} \tag{2.4}$$

$Y_{i,t}^z$ only has one random variable, $W_{i,t}$. Due to the fact that this variable is normally distributed, $Y_{i,t}^z$ also follows a Normal distribution of the form $\mathcal{N}(\mu, \sigma^2)$. The parameters can be computed as follows:
$\mu = \mathbb{E}[Y_{i,t}^z] = \mathbb{E}[\sqrt{\omega} \cdot z + \sqrt{1-\omega} \cdot W_{i,t}] = \sqrt{\omega} \cdot z + \sqrt{1-\omega} \cdot \mathbb{E}[W_{i,t}] = \sqrt{\omega} \cdot z.$
$\sigma^2 = \mathbb{V}[Y_{i,t}^z] = \mathbb{V}[\sqrt{\omega} \cdot z + \sqrt{1-\omega} \cdot W_{i,t}] = (1-\omega) \cdot \mathbb{V}[W_{i,t}] = 1-\omega.$

Note that z is not a random variable, and therefore $\mathbb{E}[z] = z$ and $\mathbb{V}[z] = 0$. Now, we can define $PD_{i,t}^z$.

Definition 2.2.2 (Probability of Default at time t conditioned to $Z_t = z$)**.** Given a threshold $s \in \mathbb{R}$, the Probability of Default at time t conditioned to some value z of Z_t ($PD_{i,t}^z$) is defined as:

$$PD_{i,t}^z := \mathbb{P}[Y_{i,t}^z < s] = \Phi\left(\frac{s - \sqrt{\omega} \cdot z}{\sqrt{1-\omega}}\right) =: f(z) \tag{2.5}$$

Where $\Phi\left(\frac{s-\sqrt{\omega}\cdot z}{\sqrt{1-\omega}}\right)$ comes from the fact that $\frac{Y_{i,t}^z - \sqrt{\omega}\cdot z}{\sqrt{1-\omega}} \sim \mathcal{N}(0,1)$, since it is the standardized version of $Y_{i,t}^z$.[7]

Only when $z < 0$, the probability of default increases, and therefore regulators are mostly interested in this particular case. It is worth remembering that $z < 0$ implies a negative state of the economy. Consequently, it can be inferred that negative states of the economy decrease the solvency of banks.
So far, we have obtained the distribution of both $PD_{i,t}$ and $PD_{i,t}^z$, but they both depend on s which depends on PD^{LR}, by the equation (2.3). In practice, PD^{LR} is unknown, and therefore it must be estimated.

2.3 From theory to practice

2.3.1 Introducing the binomial distribution

As it will be clear in the next section, in order to go from theory to practice, a basic knowledge on the binomial distribution will be required. There are many probability distributions, but in this project we only use this one, which is the sum of independent Bernoulli distributions. Hence, let us define both distributions.
The Bernoulli distribution is a fundamental probability distribution that models a random experiment with two possible outcomes: success and failure. This model only has one parameter (p) which is the probability the the event ”success” occurs. Therefore, the Probability Mass Function is:

[7] $PD_{i,t}^z$ is defined as $f(z)$ to simplify the notation.

$$P(X = x) = \begin{cases} p, & \text{if } x = 1 \\ 1 - p, & \text{if } x = 0 \end{cases}$$

And its expected value (Mean) and variance can be easily computed as follows:

$$\mathbb{E}[X] = 1 \cdot p + (1 - p) \cdot 0 = p \tag{2.6}$$

$$\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 1 \cdot p + (1 - p) \cdot 0 - p^2 = p(1 - p) \tag{2.7}$$

In our case, one borrower can either default or not, so its behaviour can be described by this distribution. Moreover, the event ”success” is going to be described as the borrower being in default, so p is going to be the probability of default (PD^{LR}). However, banks don’t only have one borrower, they have a portfolio with many borrowers and therefore the Binomial distribution is required.

The binomial distribution $B(n, p)$ extends the concept of the Bernoulli distribution to n independent Bernoulli trials, each with the same success probability p. It is actually the sum of n independent Bernoulli. It models the number of successes k in these n trials. Thus, we will be able to compute how many borrowers will default, which will be dependent on the probability of default. [8]

More rigorously, we define the binomial distribution as follows: Let’s assume that $X1, X2, ..., Xn$ are a set of independent random variables such that $X_i \sim Bernouilli(p)$, then a variable Y is a binomial distributed random variable if it is defined as the sum of X_i:

$$Y = \sum_{i=1}^{n} X_i$$

The probability mass function of the binomial distribution gives the probability of obtaining exactly k successes in n trials:

$$P(Y = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Where $\binom{n}{k}$ is the binomial coefficient, representing the number of ways to choose k successes out of n trials.

Its Mean and Variance can be easily computed using the moments of the Bernoulli as follows:

$$\mathbb{E}[Y] = \sum_{i=1}^{n} \mathbb{E}[X_i] = p + p + ... + p = n \cdot p. \tag{2.8}$$

$$\mathbb{V}[Y] = \sum_{i=1}^{n} \mathbb{V}[X_i] = p(1 - p) + p(1 - p) + ... + p(1 - p) = n \cdot p \cdot (1 - p). \tag{2.9}$$

[8] Notice that we are assuming that the borrowers are independent.

Now we are ready to deal with the real goal of this section which is dealing with the fact that PD^{LR} is not known and, thus, must be estimated.

2.3.2 Conditional default rate

In practice PD^{LR} is not known. However, the bank can evaluate the default rate at t (DR_t), that is the number of expositions found at default divided by the number of all the expositions in the portfolio, and use it as a measure of the conditioned probability of default, i.e. the Probability of default associated with a given value of the factor Z_t ($PD^z_{i,t}$).
Before defining DR_t mathematically , let us define the indicator function $D^z_{i,t}$ which is 1 when the creditworthiness conditioned to the common external factor falls below a threshold s, and 0 otherwise. Formally,

$$D^z_{i,t} = \mathbb{I}_{\{Y^z_{i,t}<s\}} \tag{2.10}$$

And since, $\mathbb{P}[Y^z_{i,t} < s] = f(z)$, $D^z_{i,t} \sim Bernoulli(f(z))$.[9]

Definition 2.3.1 (Default Rate conditioned to $Z_t = z$)**.** Let's assume the number of borrowers is N. Then, the Default Rate in the period t conditioned to $Z_t = z$ ($DR^z_{i,t}$) is defined as:

$$DR^z_{i,t} = \sum_{i=1}^{N} \frac{D^z_{i,t}}{N} \tag{2.11}$$

By definition, the sum of n independent Bernoulli's random variables with success probability p follows a binomial distribution (n, p). Therefore, $\sum_{i=1}^{N} D^z_{i,t} \sim Binomial(N, f(z))$, implying that $\mathbb{E}[D^z_{i,t}] = N \cdot f(z)$ and its variance $\mathbb{V}[D^z_{i,t}] = N \cdot f(z) \cdot (1 - f(z))$. Consequently, the expected value and variance of $DR^z_{i,t}$ can be computed as follows:

$$\mathbb{E}[DR^z_{i,t}) = \sum_{i=1}^{N} \frac{\mathbb{E}[D^z_{i,t}]}{N} = f(z) = PD^z_{i,t}$$

$$\mathbb{V}[DR^z_{i,t}] = \sum_{i=1}^{N} \frac{\mathbb{V}[D^z_{i,t}]}{N^2} = \frac{f(z) \cdot (1 - f(z))}{N}$$

For N large enough, this binomial distribution can be approximated by the following normal distribution:

$$DR^z_{i,t} \sim \mathcal{N}\left(f(z), \frac{f(z) \cdot (1 - f(z))}{N}\right) \tag{2.12}$$

Note that $DR^z_{i,t}$ is an unbiased estimator of $PD^z_{i,t}$, and its variance decreases as N increases. The sixth hypothesis is that the portfolio is infinitely grained which means that N is large enough so the effect of the idiosyncratic factor ($W_{i,t}$) is very small; therefore the variance of the estimator should be relatively small. It is also worth noticing that although $DR^z_{i,t}$ depends

[9] $D^z_{1,t}, D^z_{2,t}, ..., D^z_{N,t}$ are independent because the common factor is fixed.

on $W_{i,t}$, $PD^z_{i,t}$ only depends on $Z_t = z$. Since,

$$\lim_{N\to\infty} \mathbb{V}[DR^z_{i,t}] = 0$$

This means that as N increases the impact of the idiosyncratic factor tends to zero. And furthermore, in the limit $DR^z_{i,t} = PD^z_{i,t}$[10].

2.3.3 Unconditional default rate

Similarly to the Definition 2.3.1, the unconditional Default Rate is defined as:

Definition 2.3.2 (Unconditional Default Rate)**.** Let's assume the number of borrowers is N. Then, the Default Rate in the period t (DR_t) is defined as:

$$DR_t = \sum_{i=1}^{N} \frac{D_{i,t}}{N} \tag{2.13}$$

Where $D_{i,t}$ = $\mathbb{I}_{\{Y_{i,t}<s\}}$. It is worth noticing that $D_{1,t}, D_{2,t}, ..., D_{N,t}$ are not independent because of the effect of the common factor Z_t, and thus DR_t does not follow a binomial distribution.

By definition, $DR_{i,t}$ is a particular case of the portfolio loss relative to the portfolio's total exposure, defined as follows.

Definition 2.3.3 (Portfolio loss relative to the portfolio's total exposure)**.** For a portfolio of N borrowers, the portfolio loss relative to the portfolio's total exposure is given by:

$$L_t = L_t^{(N)} = \sum_{i=1}^{N} w_{i,t}\eta_{i,t}D_{i,t} \;\; where \; w_{i,t} = \frac{EAD_i}{\sum_{j=1}^{N} EAD_{j,t}} \tag{2.14}$$

$\eta_{i,t}$: It's also known as the Loss Given Default (LGD), and it refers to the fraction or percentage of the exposure that is expected to be lost by the lender in the event of a customer's default in the time period t.
$EAD_{i,t}$: Exposure at Default represents the amount of funds or credit exposure that a lender is exposed to when a customer defaults on their loan or credit obligation. It is the total outstanding balance or the maximum potential loss the lender may face in the event of default in the time period t.
$D_{i,t}$: The indicator function that shows whether the borrower defaults or not in the time period t.[11]
Notice that if we are dealing with an homogeneous group of N borrowers[12] and we assume that LGD is 100% (constant), then

[10] Although in practice N might not be large enough, we will assume this holds because we are under the hypothesis HP.6, which states that the portfolio is infinitely grained.

[11] Formally, $D_{i,t} = \mathbb{I}_{\{Y_{i,t}<s\}}$.

[12] $w_i = \frac{1}{N}$.

$$L_t^{(N)} = \sum_{i=1}^{N} \frac{1}{N} \cdot 1 \cdot D_{i,t} = DR_t \tag{2.15}$$

These assumptions will prevail throughout the whole paper.
Bluhm et al. (2016) proved that under these conditions:

$$L^{(N)} \xrightarrow[N\to\infty]{} \Phi\left(\frac{\Phi^{-1}(PD^{LR}) - \sqrt{\omega}\cdot Z_t}{\sqrt{1-\omega}}\right) = P(Z_t), \quad almost \;\; surely. \tag{2.16}$$

Therefore, the probability distribution of DR_t can be computed.
By definition, $DR_t \in (0,1)$, so for every $x \in (0,1)$ we have:

$$F_{PD^{LR},\omega}(x) = \mathbb{P}[DR_t \leq x] = \mathbb{P}[P(Z_t) \leq x] = \mathbb{P}\left[\Phi\left(\frac{\Phi^{-1}(PD^{LR}) - \sqrt{\omega}\cdot Z_t}{\sqrt{1-\omega}}\right) \leq x\right] =$$
$$= \mathbb{P}\left[-Z_t \leq \frac{\Phi^{-1}(x)\cdot\sqrt{1-\omega} - \Phi^{-1}(PD^{LR})}{\sqrt{(\omega)}}\right] = \Phi\left(\frac{\Phi^{-1}(x)\cdot\sqrt{1-\omega} - \Phi^{-1}(PD^{LR})}{\sqrt{\omega}}\right) \tag{2.17}$$

The last step is because of the HP.2 and the properties of the standard normal distribution. As a consequence of both, $-Z_t \sim \mathcal{N}(0,1)$.
From the relationship (2.17), we can compute the expected value and variance of DR_t.

$$\mathbb{E}[DR_t] = PD^{LR} \tag{2.18}$$

$$\mathbb{V}[DR_t] = \Phi_2(\Phi^{-1}(PD^{LR}), \Phi^{-1}(PD^{LR}); \omega) - (PD^{LR})^2 \tag{2.19}$$

Where Φ_2 is the cumulative bivariate normal distribution with correlation ω evaluated in $\left(\Phi^{-1}(PD^{LR}), \Phi^{-1}(PD^{LR})\right)$.
Notice that the equation (2.18) implies that DR_t is an unbiased estimator of the parameter PD^{LR}.
The expected value equal to PD^{LR} comes directly from the construction of $F_{PD^{LR},\omega}$, whereas the variance is not so straight forward. To compute the variance, we use the following well-known relationship:

$$\mathbb{V}[DR_t] = \mathbb{E}[DR_t^2] - (\mathbb{E}[DR_t])^2 \tag{2.20}$$

Owing to the equation (2.18), it is only necessary to compute $\mathbb{E}[DR_t^2]$, and to do so we are going to do the following workaround[13] Let's define $X_{1,t}, X_{2,t}$ as two independent standard normal random variables. And also an X_t variable as follows:

$$X_t = \frac{\Phi^{-1}(PD^{LR}) - \sqrt{\omega}\cdot Z_t}{\sqrt{1-\omega}} \tag{2.21}$$

[13]Both, the variance and the expected value, could be computed by simply following their definitions. In order to do so, we only need the probability density function, which is the derivative of $F_{PD^{LR},\omega(x)}$.

Notice that $\Phi(X_t) = P(Z_t)$, and it only depends on one random variable, and that is Z_t which is normally distributed. Therefore, $X_t \sim \mathcal{N}(\mu, \sigma^2)$, with parameters:

$$\mu = \frac{\Phi^{-1}(PD^{LR}) - \sqrt{\omega} \cdot \mathbb{E}[Z_t]}{\sqrt{1-\omega}} = \frac{\Phi^{-1}(PD^{LR}) - \sqrt{\omega}}{\sqrt{1-\omega}}$$

$$\sigma^2 = \frac{0 + (-\sqrt{\omega})^2 \mathbb{V}[Z_t]}{(\sqrt{1-\omega})^2} = \frac{\omega}{1-\omega}$$

Let's define g_{μ,σ^2} as the density of X_t. So, we can write $\mathbb{E}(DR_t^2)$ as:

$$\mathbb{E}[DR_t^2] = \mathbb{E}[P(Z_t)^2] = \mathbb{E}[\Phi(X_t)^2] = \int_{-\infty}^{\infty} \mathbb{P}[X_{1,t} \leq X_t | X_t = x] \mathbb{P}[X_{2,t} \leq X_t | X_t = x] dg_{\mu,\sigma^2}(x) =$$

$$= \mathbb{P}[X_{1,t} \leq X_t, X_{2,t} \leq X_t | X_t = x] dg_{\mu,\sigma^2}(x) = \mathbb{P}[X_{1,t} - X_t \leq 0, X_{2,t} - X_t \leq 0] \quad (2.22)$$

Thanks to applying the conditional properties of two independent variables, the problem has been simplified to computing a probability with depends on normal variables, and thus are relatively easy to express. Notice that $X_{1,t} - X_t$ and $X_{2,t} - X_t$ are normally distributed, since they are a linear combination of 2 normal variables. In fact, they follow the same distribution with parameters:

$$\mu_{X_{i,t}-X_t} = \mathbb{E}[X_{i,t} - X_t] = 0 - \mathbb{E}[X_t] = \frac{\Phi^{-1}(PD^{LR})}{\sqrt{1-\omega}}$$

$$\sigma^2_{X_{i,t}-X_t} = \mathbb{V}[X_{i,t} - X_t] = \mathbb{V}[X_{i,t}] + \mathbb{V}[X_t] = 1 + \frac{\omega}{1-\omega}$$

Note that $Cov[X_{i,t}, X_t] = 0$. Now, let us compute the correlation between $X_{1,t} - X_t$, $X_{2,t} - X_t$; to do so we will need to apply the fact that $Cov[X_{1,t} - X_{2,t}]$ is also zero by independence,

$$Corr[X_{1,t} - X_t, X_{2,t} - X_t] = \frac{Cov[X_{1,t} - X_t, X_{2,t} - X_t]}{\sqrt{\sigma^2_{X_{2,t}-X_t} \cdot \sigma^2_{X_{1,t}-X_t}}} = \frac{\omega}{1 + \frac{\omega}{1-\omega}} = \omega. \quad (2.23)$$

By means of (2.22) and (2.23) and shifting and scaling $X_{1,t} - X_t$ and $X_{2,t} - X_t$ to achieve the standard normal distribution, we conclude that $\mathbb{E}[DR_t^2] = \Phi_2(\Phi^{-1}(PD^{LR}), \Phi^{-1}(PD^{LR}); \omega)$. Moreover, applying the relationship (2.20), we get the result (2.19).
At this point we have got an unbiased estimator of PD^{LR}. Since we know its distribution function ($F_{PD^{LR},\omega}(x)$), we can now find an expression to evaluate the worst case default rate (WCDR), the maximum default rate that should be observed given a confidence level α.In other words, we will compute a quantile $q_\alpha(DR_t)$ such that

$$\mathbb{P}[DR_t \geq q_\alpha(DR_t)] = 1 - \alpha \Longleftrightarrow \mathbb{P}[DR_t \leq q_\alpha(DR_t)] = \alpha \quad (2.24)$$

By means of the cumulative distribution of DR_t (2.17), we get:

$$F_{PD^{LR},\omega}(q_\alpha(DR_t)) \quad = \quad \Phi\left(\frac{\Phi^{-1}(q_\alpha(DR_t))\cdot\sqrt{1-\omega}-\Phi^{-1}(PD^{LR})}{\sqrt{\omega}}\right) \quad = \quad \alpha \quad (2.25)$$

And solving for $q_\alpha(DR_t)$:

$$q_\alpha(DR_t) = \Phi\left(\frac{\Phi^{-1}(PD^{LR})+\sqrt{\omega}\cdot\Phi^{-1}(\alpha)}{\sqrt{1-\omega}}\right) = \Phi\left(\frac{\Phi^{-1}(PD^{LR})-\sqrt{\omega}\cdot\Phi^{-1}(1-\alpha)}{\sqrt{1-\omega}}\right) \quad (2.26)$$

In conclusion, by using the equation (2.26) and assuming that N is large enough for this to hold, we can predict any quantile in the distribution of default rates. However, it depends on the parameter PD^{LR} which is unknown, so banks need to estimate it by computing the observed default rates. Therefore, the next logical step is to find an estimation for this unkown parameter.

2.3.4 Using DR_t to estimate PD^{LR}

A reasonable estimator for PD^{LR} would be the average observed default rates observed ($\overline{DR}$) over a period of time (T). This can be expressed as:

$$\overline{DR} = \frac{1}{T}\sum_{t=1}^{T} DR_t \quad (2.27)$$

Since DR_t has the expected value and variance specified by (2.18) and (2.19), the mean and variance of $\overline{DR}$ are:

$$\mathbb{E}[\overline{DR}] = \frac{1}{T}\sum_{t=1}^{T}\mathbb{E}[DR_t] = PD^{LR} \quad (2.28)$$

$$\mathbb{V}[\overline{DR}] = \frac{1}{T^2}\sum_{t=1}^{T}\mathbb{V}[DR_t] = \frac{\mathbb{V}[DR_t]}{T} \quad (2.29)$$

Notice that in (2.29) we have used HP.4, in other words, we have assumed that DR_t are uncorrelated.[14] Thus, by means of (2.28) and (2.29) and asuming that T is large enough,

$$\overline{DR} \sim \mathcal{N}\left(PD^{LR}, \frac{\Phi_2(\Phi^{-1}(PD^{LR}),\Phi^{-1}(PD^{LR});\omega)-(PD^{LR})^2}{T}\right) \quad (2.30)$$

Now that the estimator, which can be obtained with data, has been formally presented. The next step is to substitute PD^{LR} by it in the equation 2.26, so that we obtain the estimated

[14] In real life, they are usually correlated. However this assumption was made for simplification purposes.

quantile[15]:

$$\widehat{q}_\alpha(DR_t) = \Phi\left(\frac{\Phi^{-1}(\overline{DR}) - \sqrt{\omega}\cdot\Phi^{-1}(1-\alpha)}{\sqrt{1-\omega}}\right) \tag{2.31}$$

2.3.5 A solution to correct the underestimation error

As we will see in the next chapter this estimator tends to underestimate $q_\alpha(DR_t)$, which is a problem because it means that we are facing greater risk than the one that we estimate, and therefore the reserves of money would be smaller than they should. For this reason, Casellina et al. (2023) introduced a possible solution to this problem, consisting on computing an upper bound for $\overline{DR}$, and using it as the estimator of PD^{LR}[16].
The goal is to find UP_β such that:

$$\mathbb{P}[PD^{LR} \leq UP_\beta] = \beta$$

As shown in (2.30), $\overline{DR}$ follows (asymptotically) a normal distribution with expected value PD^{LR} , and thus, we can use it to compute the upper bound (UP_β) with a confidence level $\beta \in (0,1)$. By including $\overline{DR}$ and standardizing the variable, we get:

$$\mathbb{P}\left[\frac{\overline{DR} - PD^{LR}}{\sqrt{\mathbb{V}[\overline{DR}]}} \leq \frac{\overline{DR} - UP_\beta}{\sqrt{\mathbb{V}[\overline{DR}]}}\right] = 1-\beta$$

This implies that $\Phi\left(\frac{\overline{DR}-UP_\beta}{\sqrt{\mathbb{V}[\overline{DR}]}}\right) = 1-\beta$. As a consequence, and by means of (2.29) and the symmetry of the standard normal distribution, UP_β can be expressed as:

$$UP_\beta = \overline{DR} - \Phi^{-1}(1-\beta)\cdot\sqrt{\mathbb{V}[\overline{DR}]} = \overline{DR} + \Phi^{-1}(\beta)\cdot\sqrt{\frac{\mathbb{V}[DR_t]}{T}} \tag{2.32}$$

Hence, if we substitute $\overline{DR}$ by its upper bound UP_β, the quantile expressed in (2.31) is now:

$$\widehat{q}_{\alpha,\beta}(DR_t) = \Phi\left(\frac{\Phi^{-1}\left(\overline{DR} + \Phi^{-1}(\beta)\cdot\sqrt{\frac{\mathbb{V}[DR_t]}{T}}\right) - \sqrt{\omega}\cdot\Phi^{-1}(1-\alpha)}{\sqrt{1-\omega}}\right) \tag{2.33}$$

Notice that now the quantile depends on two confidence levels α and β, so can vary the parameter the confidence level β to correct the error derived from the estimation of PD^{LR}. Basically, we will need to try to find some β such that:

$$\mathbb{P}[DR_t > \widehat{q}_{\alpha,\beta}(DR_t)] = 1-\alpha \tag{2.34}$$

[15]Notice that $\overline{DR}$ is a random variable, meaning it has variability. Whereas PD^{LR} is a constant.
[16]$\mathbb{E}[UpperBound] > \mathbb{E}[\overline{DR}] = PD^{LR}$. Therefore, it is a biased estimator of PD^{LR}.

The effects of choosing different values of β will be shown in chapter 5.

Chapter 3

Monte Carlo's theoretical foundation

Monte Carlo methods are a class of numerical techniques that rely on randomness and probability to solve complex problems. They are particularly suited to solving problems that involve uncertainty or cannot be easily solved analytically. The key idea behind Monte Carlo methods is to simulate random processes or systems repeatedly and then use the statistical properties of the results to make informed decisions or estimate values.
In the context of finance, Monte Carlo methods are employed to model and analyze a wide range of financial instruments and scenarios as it can be seen in Glasserman (2003). These methods are especially valuable for pricing options, managing risk, optimizing portfolios, and assessing the potential outcomes of investment strategies. In this project, we will simulate the behaviour of different borrowers assuming a specific PD^{LR}, and compute the estimated probability of default. This way, we will have both the real value (PD^{LR}) and the value that banks would have estimated "in real life" and therefore we can compare them and see the error of estimation.
Monte Carlo's mathematical foundation lies in probability theory, statistics, and the law of large numbers. Let's delve into the key mathematical concepts that underpin this method.

3.1 Random variables and the law of large numbers

In Monte Carlo simulations, random variables are used to model uncertain quantities or events. A random variable is a function that assigns a real number to each possible outcome of a random experiment. It is typically denoted as X and follows a probability distribution, often denoted as $f(x)$, which describes the likelihood of different values of X occurring. As seen in the previous chapter, the binomial distribution will be key in our problem. Remember that an important restriction of this distribution is the independence between experiments, in our case between the borrowers. This could be problem, since under the ASRF model are not independent due to the common factor. However, for this same reason, it was introduced in the previous chapter the notion of the indicator function conditioned to the common factor Z ($D^z_{i,t}$), which by definition eliminates this problem[1]. In fact, we defined $D^z_{i,t}$ in (2.3.1) as the sum of n n $D^z_{i,t}$, which allows us to use the binomial distribution. Notice that our parameter p

[1]Notice that now it only depends on the idiosyncratic factor, which is independent between borrowers by HP.1

is going to be $f(z)$, which is the probability of default at the period t conditioned to $Z_t = z$.
Another important concept is the law of large numbers. This Law is the one that assures that by using Monte Carlo we are actually getting closer to the solution. It is a fundamental theorem in probability theory and statistics and describes the behavior of sample averages as the size of the sample or the number of trials increases. It can be stated as follows:

Definition 3.1.1 (Law of Large Numbers)**.** Let $X_1, X_2, X_3, \ldots, X_n$ be a sequence of independent and identically distributed (i.i.d.) random variables with mean μ, and let $\bar{X}_n$ be the sample average of the first n observations:

$$\bar{X}_n = \frac{X_1 + X_2 + X_3 + \ldots + X_n}{n}$$

Then,:

$$\lim_{n\to\infty} \bar{X}_n = \mu$$

In our particular case, we are using Monte Carlo to estimate $q_\alpha(\{DR_t\})$. Since, $\mathbb{E}[DR_t^z] = DR_t$ and we can compute DR_t^z by means of the binomial distribution, we can get a series of default rates which allows us to obtain a series of n quantiles by means of the equation (2.31)[2] . Notice that we are working with an unbiased estimator of PD^{LR}, so our quantile estimator is also unbiased.
Brandimarte (2014) shows that Monte Carlo has an error of order $\mathcal{O}(\frac{1}{\sqrt{n}})$ and, therefore, it requires a large number of simulations. For all the results in this paper, the number of simulations is 10 millions, the same number that Casellina et al. (2023) used. They also applied the variance reduction technique of importance of sampling, so we will do the same. Nevertheless, we have also used another technique which has produced better results, meaning that has taken us closer to the actual PD^{LR}. Both techniques will be explained in the next section.

3.2 Variance reduction

As previously explained, Monte Carlo methods are rather slow, and as a consequence, it is useful to apply Variace-Reduction Techniques which are used to improve the precision of the simulation-based estimates. All these methods try to get us closer to the actual solution with less replicates than the classical Monte Carlo, thus, their goal is to increase the speed at which we find a good enough solution[3].
Although useful, these techniques tend to be mathematical and statically complex[4] and its ability to improve the method varies case by case. That is to say, that even though one method might the best in one scenario it might be the worst in another. Thus, it is always interesting to try different variance-reduction techniques. In this paper we will apply two common ones.

[2]Each quantile is computed with a different estimated DR_t, obtained from the same PD^{LR}.
[3]The "good enough" solution will vary case by case. Usually, it depends on the error we are willing to accept.
[4]See Bolder (2018)

Antithetic variate

This is one of the simplest methods from the mathematical point of view, but it sometimes provides better results than more complex methods. The Antithetic variate technique is based on the idea of generating at each path of the simulation, two negative correlated variables[5]. In our case, we will use the fact that if $Z \sim \mathcal{N}(0,1)$, then $Z * (-1) \sim \mathcal{N}(0,1)$.[6]. Thanks to this property, we could use the same simulation to have 2 different scenarios, just by multiplying the result by 1 and -1.
In our case, we need to simulate the common factor (Z_t) which follows a standard normal distribution, in order to compute the probability of default conditioned to this common factor ($f(z)$). This method allows us to determine two different probabilities at each step and then taking the average of both. Here's an example:
Let's say that we get, through simulation, the value $Z_t = 0.2$, then we would compute two different probabilities by means of (2.5):

$$f_+ = f(z = 0.2) = \Phi\left(\frac{s - \sqrt{\omega} \cdot 0.2}{\sqrt{1-\omega}}\right)$$

$$f_- = f(z = -0.2) = \Phi\left(\frac{s + \sqrt{\omega} \cdot 0.2}{\sqrt{1-\omega}}\right)$$

Finally, we get our estimate ($\widehat{f(z)}$) as the average of both values. This way, we will be able to find a more representative probability at each simulation, notice how the probabilities that we obtain by using this method will tend to be towards the center, meaning that it disregards rarer values.
It is fairly easy to see how this method does not only reduces the samples of normal distributions, but also reduces the variance for every path. Notice that the variance of our estimate is:

$$\mathbb{V}[\widehat{f(z)}] = \mathbb{V}[0.5 \cdot (f_+ + f_-)] = \frac{\mathbb{V}[f_+] + \mathbb{V}[f_-] + 2Cov[f_-, f_+]}{4}$$

By definition f_+ and f_- are negatively correlated and, therefore, their covariance is negative. Plus, $\mathbb{V}[f_-] = \mathbb{V}[f_+]$. Thus,

$$\mathbb{V}[\widehat{f(z)}] = \frac{\mathbb{V}[f_+] + \mathbb{V}[f_-] + 2Cov[f_-, f_+]}{4} < \frac{\mathbb{V}[f_+] + \mathbb{V}[f_-]}{4} = \frac{\mathbb{V}[f_+]}{2} < \mathbb{V}[f_+]$$

In conclusion, by using this method the variance of our estimate is now smaller and we should get a better approximation to the solution with less simulations.

Importance of sampling

This is a more complex method based on the idea of tweaking the probability distribution in such a way that less likely outcome becomes more probable as it can be seen in Hammersley

[5]See Hammersley and Morton (1956) and Kleijnen et al. (2010)
[6]This follows from the properties of the expected value and the variance.

(2013). In order to understand the concept, it is useful to remember that Monte Carlo is based on the idea of averages. In fact, we always try to compute the following expectation:

$$\mathbb{E}_p[f(x)] = \int p(x)f(x)dx \tag{3.1}$$

where $p(x)$ is a probability density function and $f(x)$ the function of which we are trying to compute the expectation. The goal is to approximate this by computing the average. To do so, we take random samples of x, taking into account that $x \sim p(x)$ and compute the average of $f(x)$ evaluated at those points. Mathematically,

$$\mathbb{E}_p[f(x)] \approx \frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

By the central limit theorem, our estimate can be approximated by a normal distribution, when N is large enough, with $\mu = \mathbb{E}_p[f(x)]$ and $\sigma^2 = \frac{1}{N}\mathbb{V}_p[f(x)]$.
Notice how if $f(x)$ is bigger in values of x with lower probabilities, it will require many simulations in order to get a good enough estimation, since the values that would have a greater impact on the expectation are less likely to be observed. Bearing this idea in mind, importance of sampling tweaks this probability so we can get a better approximation faster. We accomplish this by introducing a new probability function $(h(x))$[7], as follows:

$$\mathbb{E}_p[f(x)] = \int \frac{h(x)}{h(x)} p(x)f(x)dx = \int h(x)\left(\frac{p(x)}{h(x)}f(x)\right)dx = \mathbb{E}_h\left[\frac{p(x)}{h(x)}f(x)\right] \tag{3.2}$$

The ratio $\frac{p(x)}{h(x)}$ is often referred as the Radon-Nikodym derivative or the likelihood ratio and its function is to rescale f(x). Now, our Monte Carlo estimator is:

$$\widehat{\mathbb{E}[f(x)]} = \frac{1}{N}\sum_{i=1}^{N} f(x_i)\frac{p(x_i)}{h(x_i)}$$

Although $h(x)$ does not need to meet any special requirements, besides being a probability density function, its choice will determine the effectiveness of this method. A good $h(x)$ will allow us to get a good enough estimation faster, whereas a bad choice will produce the contrary effect. Usually we choose the $h(x)$ that minimizes the variance of the Monte Carlo estimator, but this is not trivial.
In our problem, the probability that we will try to change is the probability of default and to do so, we will use the solution proposed by Glasserman and Li (2005), known as the Exponential twisting or Esscher transform. They have applied this to the computation of capital for credit risk in the multifactorial Merton's model, which is defined as follows:

Definition 3.2.1. (Normal Copula Model) Let's assume N obligors, $c_k \in [0, 1]$, the loss given default of the kth obligor, and $Y_k = \mathbb{I}_{\{\eta_k > x_k\}}$ as the default indicator of the kth obligor ($Y_k = 1$ if defaults, 0 otherwise). Then, the total loss is:

[7]h(x) just needs to meet the standard definition of statistical density function.

$L = c_1 \cdot Y_1 + c_2 \cdot Y_2 + ... + c_N \cdot Y_N$

Where $x_k = \Phi(1 - PD^{LR})$ and $\eta_k = a_{k1} \cdot Z_1 + ... + a_{kd} \cdot Z_{kd} + b_k \epsilon_k$, in which
- $Z_1, ..., Z_d$ are systematic risk factors.
- $Z_i \sim \mathcal{N}(0, 1)$
- ϵ_k is the idiosyncratic risk of the kth obligor, also normally distributed with mean 0 and variance 1.
- $a_{k1}, ..., a_{kd}$ are the factor loadings for the k-th obligor, $a_{k1}^2 + ... + a_{kd}^2 \leq 1$.
- $b_k = \sqrt{1 - (a_{k1}^2 + ... + a_{kd}^2)}$ so that, $x_k \sim \mathcal{N}(0, 1)$.

Under this definition, Glasserman (2005) proposed the following transformation on the probability of default conditioned on the common factor ($p_k(z)$):

$$p_k(\theta, z) = \frac{p_k(z) e^{\theta c_k}}{1 + p_k(z)(e^{\theta c_k} - 1)} \tag{3.3}$$

Where $\theta \in \mathbb{R}$ is a parameter that can be computed and determines the transformation. If $\theta > 0$, the probability of default will be incremented, while if it is negative it will be decreased. In case $\theta = 0$, the probability of default remains unchanged. Consequently, we will be interested in $\theta > 0$.
This Normal Copula Model might seem a little bit different than our model, but notice how we get the same model if instead of d systematic risk factors, we have only one and we establish the following relationships:
- $a_{k1} = \sqrt{\omega}$
- $b_k = \sqrt{1 - \omega}$
- $\epsilon_k = W_{i,t}$
- $Z_1 = Z_{t,k}$
- $c_1, ...c_k$ are 100%.
The state of default is also defined somewhat different, but they are equivalent. This comes from the fact that the normal distribution is symmetric, and thus, $x_k = \Phi^{-1}(1 - PD^{LR}) = -\Phi^{-1}(PD^{LR}) = -s$.
In conclusion, we are dealing with a particular case of the Normal Copula Model which is the one-factor Merton model, and therefore it makes sense to use the transformation (3.3). However, it is worth noticing that since the loss given default in our model is 100% our probability will be the same for the kth obligors[8] . Then,

$$p(\theta, z) = \frac{f(z) e^{\theta}}{1 + f(z)(e^{\theta} - 1)} \tag{3.4}$$

Now our goal is to find θ. To do so, we will use the fact that coming from the general definition of the Radon-Nikodym derivative we get that the likelihood ratio in our model is (See Bolder (2018)):[9]

[8] We will use $f(z)\ \forall k$. We are using the notation of our model, meaning that our indicator function is $D_{i,t}^z$ and our threshold is s.

[9] Note how this is the classical likelihood ratio of a binomial distributed sample, taking into account that our probability now is divided by $h(.) = p(\theta, z)$.

$$\Pi_{i=1}^{N}\left(\frac{f(z)}{p(\theta,z)}\right)^{D_{i,t}}\left(\frac{1-f(z)}{1-p(\theta,z)}\right)^{1-D_{i,t}} \tag{3.5}$$

Where $D_{i,t}$ is our default indicator defined as in the equation (2.10). By replacing $p(\theta, z)$ by its definition(3.3) in the relationship (3.5), we get the following simplified version:

$$\Pi_{i=1}^{N}\left(\frac{f(z)}{p(\theta,z)}\right)^{D_{i,t}}\left(\frac{1-f(z)}{1-p(\theta,z)}\right)^{1-D_{i,t}} \quad = \quad \Pi_{i=1}^{N}\left(e^{\theta}\right)^{-D_{i,t}}\left(1+f(z)\cdot(e^{\theta}-1)\right) \tag{3.6}$$

A useful step towards simplfying this expresion, is taking the natural logarithm:

$$ln\left(\Pi_{i=1}^{N}\left(e^{\theta}\right)^{-D_{i,t}}\left(1+f(z)\cdot(e^{\theta}-1)\right)\right) = -\theta\sum_{i=1}^{N}D_{i,t}+\sum_{i=1}^{N}ln\left(1+f(z)\cdot(e^{\theta}-1)\right) = -\theta L+\psi(\theta)$$

Thus,

$$\Pi_{i=1}^{N}\left(e^{\theta}\right)^{-D_{i,t}}\left(1+f(z)\cdot(e^{\theta}-1\right) = e^{-\theta L+\psi(\theta)}$$

Notice how $\psi(\theta)$ is the cumulant generating function (CGL) of L[10] and therefore, this transformation of the default probabilities is equivalent to directly ajusting the default-loss density. As seen in Glasserman and Li (2005), in order to reduce the variance of our estimator we can find our θ by solving the following equation:

$$\frac{\partial}{\partial\theta}\psi(\theta,z) = \sum_{i=1}^{N}p(\theta,z) = s$$

Since $p(\theta, z)$ does not depend on i,

$$N\cdot p(\theta,z) = s \longrightarrow \theta_s(z) = ln\left(\frac{s(1-f(z))}{f(z)(N-s)}\right)$$

$\theta_s(z)$ is actually a function which depends on the value of the common factor, that means that it will change at every simulation. And in fact, since only the positive values increase the probability of default, our $\theta_s^*(z) = max(0, \theta_s(z))$.
To sum up, if $\theta_s(z) \leq 0$ the default probability will stay the same[11], but if $\theta_s(z) > 0$ the default probability will be tweaked by means of the expression (3.4).

[10] $\psi(\theta) = ln(\mathbb{E}[e^{\theta L}])$.
[11] This means that $p(\theta, z) = f(z)$.

Chapter 4

Application and results

By now, every piece of theory required to tackle the problem ahead has been explained. Now let's focus on the practice. Our main goal is to estimate the error that banks are making by assuming that the probability of default that they are estimating is the actual PD^{LR}. In order to do so, we will assume some PD^{LR} and produce some samples[1]. From this samples, we will estimate the probability of default, just as banks would do in real life, and compare both the quantile obtained with the real value and the estimated one. We will realize that by using the estimation we are underestimating the error, so we will try to solve this problem by estimating an upper bound.

4.1 Implementation in R

As mentioned throughout the paper, different common factors will be simulated in order to estimate the probability of default. Banks cannot use a period T smaller than 5 to estimate their probability, so we will take T=5. Moreover, we will repeat the experiment 10 million times. Both the time period and the number of experiments were set by Casellina et al. (2023) and since we're trying to replicate their results, we must use the same parameters. Thus, we will assume a correlation $\omega = 0.24$,the maximum value provided by the regulation, and that the number of borrowers is $N = 50,000$.
The algorithm to obtain the Worst Case Default Ratio (WCDR) is:

Algorithm 1 Classical Monte Carlo

1. Generate T common factors Z, from a $\mathcal{N}(0, 1)$.
2. Compute the probability $f(z)$ with the expression (2.5).
3. Compute T default ratios (DR^z) using the binomial distribution with $p = f(z)$.
4. Take the average of the default ratios.
5. Compute $WCDR_i$ using the expression (2.31)
6. Repeat these steps 10 Million times and compute the average $WCDR$.

Source: own elaboration.

[1] We will do 10 millions replicates.

Notice how in this algorithm we have not applied any variance reduction technique. However, both the importance of sampling and the antithetic variate technique affect only the step number 2 of the algorithm. In order to implement them, $f(z)$ just needs to be changed using the steps specified in the section 4.1.3 Variance Reduction for each method.[2]

4.2 Results

Our main goal is to see the difference between computing the WCDR through the actual $PD^{LR}(q_\alpha(\{DR_t\}))$ and with the estimation of this probability $(\mathbb{E}[\widehat{q}_\alpha(\{DR_t\})]$, as banks do in real life. The following table shows the results for different probabilities of default (PD^{LR}) and different confidence levels α.[3]

Table 4.1: α-quantiles of DR_t (WCDR)

			$--PD^{LR}--$			
α	Quantile	Variance Reduction	0.30%	1.00%	5.00%	10.00%
	$q_{99\%}(\{DR_t\})$	-	3.25%	8.67%	28.11%	43.54%
99.00%	$\mathbb{E}[\widehat{q}_{99\%}(\{DR_t\})]$	I.Sampling	3.04%	8.17%	26.96%	42.15%
	$\mathbb{E}[\widehat{q}_{99\%}(\{DR_t\})]$	Antithetic	3.16%	8.48%	27.84%	43.31%
	$q_{99.5\%}(\{DR_t\})$	-	4.42%	11.10%	33.02%	49.10%
99.50%	$\mathbb{E}[\widehat{q}_{99.5\%}(\{DR_t\})]$	I.Sampling	4.07%	10.38%	31.58%	47.48%
	$\mathbb{E}[\widehat{q}_{99.5\%}(\{DR_t\})]$	Antithetic	4.26%	10.83%	32.68%	48.84%
99.90%	$\mathbb{E}[\widehat{q}_{99.9\%}(\{DR_t\})]$	I.Sampling	7.09%	16.23%	41.99%	58.52%

Source: own elaboration.

$q_\alpha(\{DR_t\})$ is the actual quantile and as we can see in the table, the estimated quantiles are smaller, and therefore banks are underestimating the estimation error.

Remember that banks are using their own estimation which, in case of the Importance of Sampling, for an actual (unknown) probability of default 0.3% and $\alpha = 99\%$ would be 3.04% when the real value would be 3.25%. That is to say, that they are assuming that the quantile is 0.21% points smaller than it actually is and this has a direct impact on the reserves.

Notice that the same can be observed with the different PD^{LR} and the absolute difference increases as the long run probability of default gets bigger.

The actual error that banks are making is hard to know, since we don't actually know the real probability of default (PD^{LR}) and, as can be seen in the table, there is a big difference in our simulation, just by applying a different variance reduction technique. Observe how the estimated quantile applying the Antithetic Variates is closer to the actual value than the one

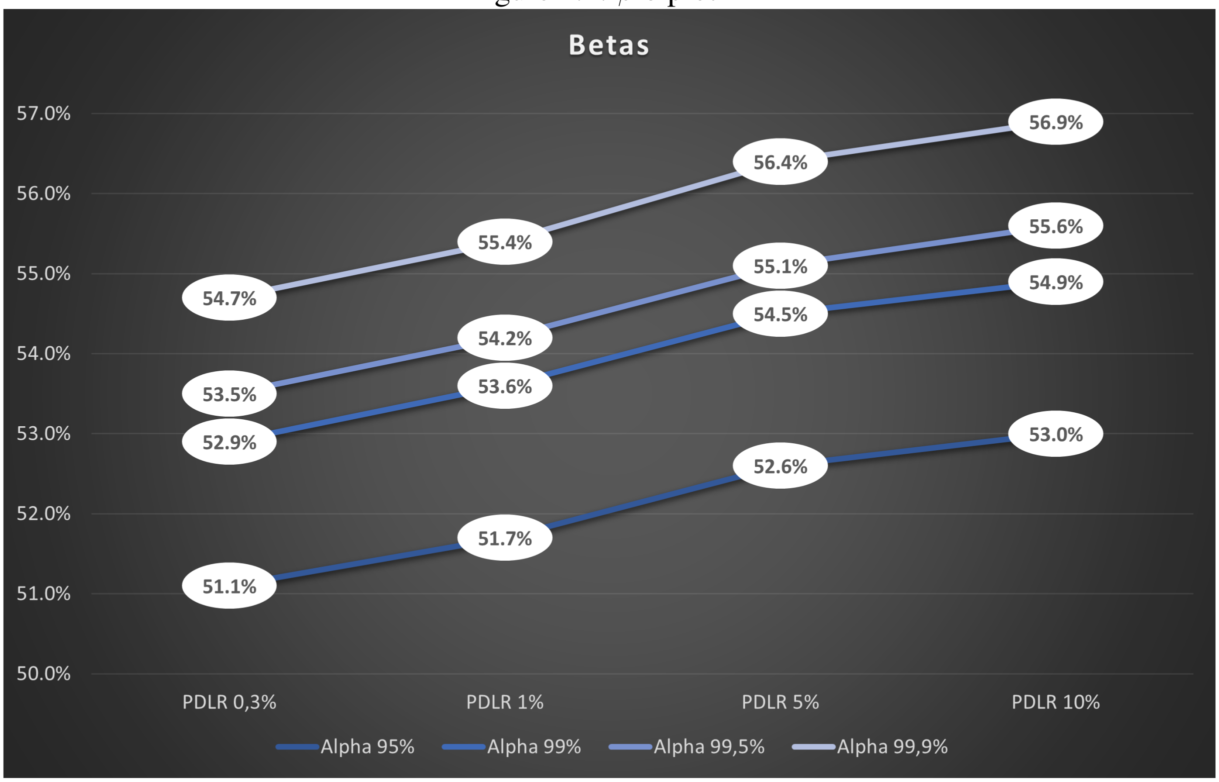
[2]For the Importance of Sampling technique applying the Esscher transform and for the Antithetic variate taking the average of f_- and f_+.

[3]These values have been set based on the work from Casellina et al. (2023)

implementing importance of sampling. Although the later is a much more complex technique.

Basel Committee on Banking Supervision (BCBS) already knows this, and that's the reason why they ask banks to set the confidence level to 99.9%[4], instead of 99% or 99.5%. As it can be seen in the table, this change on α effectively corrects the underestimation. It is also worth noticing that as the actual long run probability gets bigger the difference between $q_{99\%}(\{DR_t\}))$ and $\mathbb{E}[\widehat{q}_{99.9\%}(\{DR_t\})]$ gets smaller.
Even though setting a higher confidence level works, Casellina et al. (2023) proposed introducing an upper bound on the estimated default rate to compute the quantile (2.33)[5]. This solution requires us to determine a value β that guarantees that the default rate is bigger than the quantile with probability $1-\alpha$ (2.34). The goal is to find the smallest β that corrects the estimation error.
In the following plot it can be seen the value of β requiered for different values of PD^{LR} and α.

Figure 4.1: $\beta's$ plot



Source: own elaboration.
As α increases, the value of β necessary to fix the error due to estimating the parameter also increases. Similarly, the bigger the value of the parameter PD^{LR}, the bigger the value of β. However, all values are within the $50-57\%$ range and therefore regulators should be able to establish a specific β in this scenario, even though the real value of PD^{LR} is unknown. In

[4]This is for market risk, for credit risk is 99.99%

[5]This complies with the European Banking Authority(EBA) requirement of computing the MoC taking into account dispersion of the estimator.

most cases, they can be sure that the long probability of default will be smaller than a specific value; and therefore they could use this value to the β required.

Chapter 5

Conclusions

To compute risk measures such as Value-at-Risk(VaR) and Expected Shortfall(ES), it is necessary to estimate the probability of default. However, banks are now assuming that this parameter is known and, therefore, not taking into account the uncertainty that comes from the estimation of this parameter, implying a possible underestimation of the required capital. To solve this, The Basel Committee on Banking Supervision (BCBS) set a Margin of Conservatism (MoC)[1] that should cover this underestimation. In this paper, we check whether this MoC is enough and study an alternative solution to compute the VaR proposed by Casellina et al. (2023). We found out that banks are actually underestimating the required capital, but this underestimation depends on the real probability of default PD^{LR} which must be estimated. Therefore, we cannot know beforehand exactly by how much will they underestimate the VaR. Nonetheless, we found that by building an upper bound (2.32) that takes into account PD^{LR} and α[2] regulators should be able to set a parameter β which overcomes the underestimation error, even though it is not constant. In fact, for the values that we have used $\beta \in (0.50, 0.57)$. As Casellina et al. (2021) commented, this has been developed being coherent with the ASRF framework without introducing any extra hypotheses or other elements such as the prior distributions or other parameters which, having to be estimated, would introduce another source of estimation error.

It has also been found that the MoC actually covers the underestimation error, but it makes banks reserve substantially more capital that it would be necessary, especially when the real PD^{LR} is small. However, by building the upper bound we could find a less conservative solution and not force banks to reserve way more capital than what they actually need.

In the beginning, the main goal of this project was just to reproduce the results obtained by Casellina, but we have also introduced a variation on the Monte Carlo simulation which reduced the number of simulations required. This is the Antithetic variate variance reduction technique which is a less complex technique than the one use by Casellina, but has produced better results, meaning an estimation closer to the real value with the same number of simulations.

[1]They set a higher confidence level. For example, for market risk they set a confidence level α of 99.9% instead of 99%

[2]The confidence level.

Appendix A

R code

Click here to see the code: R code[1]

[1] If you can't click on it, go to the next url: `https://github.com/AibersonVentura/TFM2024`

Bibliography

Basel Committee, B.S., 2005. Basel ii: International convergence of capital measurement and capital standards: A revised framework. Bank for International Settlements, Basel .

Bluhm, C., Overbeck, L., Wagner, C., 2016. Introduction to credit risk modeling. Crc Press.

Bolder, D.J., 2018. Credit-Risk Modelling. Springer.

Brandimarte, P., 2014. Handbook in Monte Carlo simulation: applications in financial engineering, risk management, and economics. John Wiley & Sons.

Casellina, S., Landini, S., Uberti, M., 2021. The estimation risk and the irb supervisory formula. European Banking Authority Research Paper .

Casellina, S., Landini, S., Uberti, M., 2023. Credit risk measures and the estimation error in the asrf model under the basel ii irb approach. Communications in Nonlinear Science and Numerical Simulation 118, 106977.

EBA, E.B.A., 2017. Guidelines on pd estimation. in: Lgd estimation and the treatment of defaulted exposures. .

Glasserman, P., 2003. Monte Carlo Methods in Financial Engineering. Springer, New York, NY, USA.

Glasserman, P., 2005. Measuring marginal risk contributions in credit portfolios. FDIC Center for Financial Research Working Paper .

Glasserman, P., Li, J., 2005. Importance sampling for portfolio credit risk. Management Science 51, 1643–1656.

Hammersley, J., 2013. Monte carlo methods. Springer Science & Business Media.

Hammersley, J.M., Morton, K.W., 1956. A new monte carlo technique: antithetic variates. Mathematical Proceedings of the Cambridge Philosophical Society 52, 449–475.

Huang, X., Oosterlee, C., 2007. Computation of risk contribution in the vasicek portfolio credit loss model. PAMM 7, 1081103 – 1081104.

Kleijnen, J., Ridder, A., Rubinstein, R., 2010. Variance reduction techniques in monte carlo methods .

Osmundsen, K., 2018. Using expected shortfall for credit risk regulation. Journal of International Financial Markets, Institutions and Money 57.