The Default Risk Charge approach to regulatory risk measurement processes

Bonollo, Michele - IMT Lucca
michele.bonollo@imtlucca.it

Di Persio, Luca - Dept. Computer Science - UniVr
luca.dipersio@univr.it

Prezioso, Luca - UniTn-UniVr-Paris Diderot
luca.prezioso@unitn.it

Abstract

In the present paper we consider the Default Risk Charge (DRC) measure as an effective alternative to the Incremental Risk charge (IRC) one, proposing its implementation by a quasi exhaustive-heuristic algorithm to determine the minimum capital requested to a bank facing the market risk associated to portfolios based on assets emitted by several financial agents. While most of the banks use the Monte Carlo simulation approach and the empirical quantile to estimate this risk measure, we provide new computational approaches, exhaustive or heuristic, currently becoming feasible because of both the new regulation and to the *high speed - low cost* technology available nowadays. Concrete algorithms and numerical examples are provided to illustrate the effectiveness of the proposed techniques.

Keywords: Default Risk Charge, Incremental Risk Charge, Quasi Exhaustive-Heuristic Algorithms

AMS classification: 91G40, 91G60, 91B70, 65C20, 60H30, 60H35

1 Introduction

The financial crisis obliged the authorities to improve in a dramatic context the banking regulation about the risk management processes and the related capital requirements. The new rules along with the awareness about the weakness of the current practices implied a new exciting research era in the risk management field, covering the whole *end to end* process, from new previously *forgotten* risk sources (systemic risk, liquidity risk, etc) to the attempt to quantify the *model risk*, to the algorithmic effort in order to get faster, granular, reliable risk measures, according to the best reporting standard, such as *auditability, drill down* features and so on. In this wide scenario, one of the challenging tasks is to capture the credit risk of the financial instruments, hence removing the classical old boundary where only for the banking book instruments, namely mortgages and loans, the credit risk was measured, while for the financial instruments in the trading book the credit risk calculation was not requested. In fact, it was

prescribed only within the so called *Basel 2.5* regulation established starting from 2011, see, e.g., [5], for further details.

The regulation path to this unified view has not yet been completed, but a new relevant step has been developed in the last 2 years with the new Basel framework for the market risk. The new regulation and the new capital charges for the default risk of the trading book portfolio permit to evaluate some alternative computational tools, that can be compared to the usual Monte Carlo approach (MC from now on) used in this area. The paper is organized as follows:

- Section 2 describes in a self consistent style the financial context and the main objective of the work;

- Section 3 introduces the mathematical framework underlying the financial settings,

- Section 4, besides containing the outline of the usual model for the default risk calculation, describes the proposal of a quasi exhaustive-heuristic algorithm along with its set-up for real cases.

2 The financial context and goal of the work

Since the risk management is a vast field, enriched by a large amount of different applications, spanning from pure practitioners' ones to more theoretically oriented subjects, we have provided some references to avoid any possible misunderstanding concerning both the main goal characterizing the present work and the general context within which it has been considered.

First of all, let us underline that we can distinguish two main different paths in the risk management history. The first one is the scientific risk management, namely the whole set of models, and mathematical techniques developed by both the scientific and the professional community. Such models are often used by banks, according to internal scrutiny procedures mainly aiming at adapting them to real scenarios and contingent decisions.

The second path is constituted by the so called *regulatory risk management*, which is nothing but the set of rules the banks are requested to apply to measure their risks. The latter point is strictly linked to the banks' obligation to have enough capital to prevent them from huge losses. In this direction, a fundamental regulatory framework is the one represented by the Basel committee on Banking Supervision, or BCBS for brevity. BCBS outlines and updates for each topic the proper framework, then each country adapts them to build their actual regulation, taking care to respect some general not negotiable financial constraints.

The two paths meet and interact, very frequently. Typically, new techniques and models are accepted in the general regulation framework, provided there exists a robust awareness about them.

A milestone example is represented by the competition between the Value at Risk (VaR) and Expected Shortfall (ES). The VaR was adopted as the official market risk measure in 1996 by BCBS, see, e.g., [3]. Then, the academic community pointed out its drawbacks, such as the lack of sub-additivity property, see, e.g., [1, 2]. Nevertheless, such risk measure has been adopted so far and is going to be revised within the incoming *Fundamental Review of Trading Book*

(FRTB) framework forecasted to enter in force only from 2022, when the VaR will be replaced by the ES, see [7], for further details. In this scenario, our work has been mainly focused to consider the credit risk in the trading book, namely to consider the *risk* associated to a portfolio of financial instruments held by the banks, such as bonds and equities. Even if the credit risk field is a very large area, it is possible to split it in two main research areas. The first one is represented by the development and analysis of suitable *default models* for each counterparty, along with the estimation of related default probabilities. The second major theme is the so called *portfolio credit risk* which, roughly speaking, aims to optimize the way to collect each debtor's risk to obtain a consistent risk measure for the whole portfolio. The latter implies a difficult point to be solved, namely how to properly infer the default correlations among different debtors that have not yet been observed at the calculation date. Let us recall that, concerning the default probabilities evaluation, the most celebrated contributions are the structural models by Merton, see [22], while, to what concerns the portfolio credit risk, a fundamental seminal work dates back to the Vasiceck contributions, see [27]. The banking industry has tried to exploit these first scientific contributions, working on their mathematical peculiarities, to develop more effective calculation processes, as in the case of the *binomial-based model* by *Credit Suisse* and the *Credit Metrics model*, for further details see [13], resp. [24]. The embedding of previously mentioned developments within the *Credit Risk* general framework started in in 2006, when the Basel 2 regulation admitted the statistical models to measure the credit risk of the portfolio. Until that date, only the standard models where allowed, which implied to consider a set of grids of coefficients, each of which applied to a different exposure category. Along this regulatory line, the statistical model prescribed by the BCBS has been the Gordy one (2003), also known as the *Asymptotic Single Risk Factor* model (ASRF). The ASRF model is characterized as follows: each debtor (i) has a behavior Y_i defined by

$$Y_i = b_i \cdot X + \sqrt{q - b_i^2} \cdot w_i \,,$$

where $w_i, X \sim N(0,1)$, with $\rho(X, w_i) = 0$. In particular, X is the single systematic risk factor and all the debtors depend on it by the *factor loading b*, while the specific features are summarized in the independent Brownian-type noise represented by w_i. The counterparty defaults if Y_i is below a threshold K, given by $K = \Phi^{-1}(P)$, P being the default probability of (i) that is estimated by other statistical models. We underline that ASRF model is very appealing for the regulator, in fact, if the credit portfolio has many positions with exposures amount E_i and it holds the perfect granularity property, i.e.

$$\frac{E_i}{\sum_{j=0}^{n} E_j} \to 0 \quad \forall i \,,$$

then the VaR of the whole portfolio can be analytically obtained by summing up the risk contributions coming from each position. It is worth to mention that, in real markets, the perfect granularity property does not hold perfectly, nevertheless if the credit portfolio is very large and not much concentrated, then the analytical VaR formula represents a good approximation. We have to consider that typical banks' positions belong to two broad categories, namely:

the *banking book* and the *trading book*. While the banking book mostly consists of classical credit products, such as loans, mortgages and so on, the trading one is mainly constituted by structured financial instruments, such, e.g., bonds, equities, derivatives. For the sake of simplicity, we do not analyze the accounting perspective that allows to classify also some financial portfolio in the banking book category, if there is not a trading purpose for those positions. One of the most important weaknesses of the Basel 2 regulation exactly lies in the *banking vs trading book classification* challenge. More explicitly, the default risk has to be measured within the regulatory framework only for the banking book portfolio, not for the trading book. As an example we consider a portfolio of plain vanilla bond. Only the interest rate risk, which is referred as the *generic risk* in the BCBS language, and the spread risk, or *specific risk* from the regulation point of view, are captured by a 10 day 99% VaR, without any measurement and capital constraint assignment for the portfolio default risk.

From the 2007-2008 crisis we learned that also the big banking institutions and large corporates can fail, implying a huge amount of losses in the trading book portfolio, also because of possible *contagion phenomena*, see, e.g., [8] and references therein. These type of losses can not be absorbed by the bank capital, at least if no provision for the risk has been previously stated. To solve this gap an updated regulatory directive, namely the *Basel 2.5* one, has been developed, see [6] and [5] for further details, allowing, in particular, for a new risk parameter called *Incremental Risk Charge* (IRC).

This new risk measure is a 99.9% VaR with 1 year horizon. It takes into account both default risk and migration risk, or *down grade*, and it has been mainly prescribed for bonds-type instruments. The Basel 2.5 regulation came into force in 2012, accompanied by a lot of criticisms because of its overreaction properties that obliged banks to immobilize huge amount of their capital. In particular, the IRC measure is an example of *risks double counting*. In fact, the migration risk captured by the 99.9% 1 year VaR clearly overlaps the *old* classical risk measure for the market risk, i.e. the 99% 10 days VaR. Such an issue generates too high risk figures, with banks unable to efficiently allocate the requested capital. Therefore, BCBS published an updated version of the previous regulatory rules, providing the *fundamental review of the trading book* (FRTB), see, e.g., [7], which, even if becoming into effect just starting from 2022, has immediately pushed banks to be compliant with it, by applying its the new rules in advance within their reporting and limits system as well as to try to develop properly adapted internal models. Therefore, the FRTB has allowed to overcome some of the previously present *financial misunderstandings*. The IRC has been replaced by the *Default Risk Charge* (DRC).

The main features of DRC are:

- It is still a 99.9% 1 year VaR

- Only default risk is considered, while migration risk is removed

- Equity style positions have to be considered

- The default model must be a 2 *systematic factors model*, to overcome the drawbacks of ASRF. In fact, since trading book often consists of a relative small number of concentrated positions, namely just dozens or hundreds of debtors, instead of set many thousands, the ASRF proxy could provide too rough estimates.

- The default correlations must be jointly estimated, hence taking into consideration both equity prices and spread movements

An excellent review of the FRTB regulation can be found in [20], while the new DRC challenges are clearly explained in [25]. Within this scenario, we do not aim to suggest what systematic risk factors have to be selected, how to estimate the default correlations, how to map illiquid instruments, or what is the proper proxy to be considered. Instead, in the global DRC calculation work-flow, we try to innovate in the final quantile calculation procedure. In other words, once the model have been stated, the positions have been classified and we know the joint default probability of all the debtors, we want to compare the classical Montecarlo approach with deterministic exhaustive or near exhaustive procedures.

It is worth to observe that the more the quantile level is extreme, 99.9% for DRC, the more the Montecarlo empirical quantile can suffer of high variance estimation error, as outlined, e.g., in [14].

Then a huge number, such as 10 or 100 millions, of heavy simulations are run by the banks to get convergence of the empirical estimator, implying a high time consuming procedure which is also not so easy to set-up.

At the best of our knowledge, our analysis innovates the already present literature on the subject, showing a new proposal that we are confident can be a very promising alternative to the existing ones.

3 The mathematical setting
Model for the IRC/DRC measure

To focus on the statistical and algorithmic problem, we skip some of the several technical details of the regulation, and we simplify a bit the complex model. Roughly speaking, most of the model for IRC are *structural* models, where the default event of each issuer is related to some *background* risk factors. Moving from *IRC* to *DRC*, the model for the portfolio losses could remain the same, only the events to be considered are changing.

For a general overview of the structural model *á la* Merton, see the seminal paper in [22].

For the practical implementation in the industry, the benchmark model is the *CreditMetrics* model, see [24].

For the sake of simplicity we focus now on the most common model in the banking practice, avoiding too many theoretical definitions and preliminaries.

NOTATION

- J: the number of issuers (of bonds, equities, ...) in the trading portfolio. Usually for medium banks J could be of some dozens, more than one hundred only for very large banks. Moreover, we have a *concentration* effect, e.g. with the top 10 issuers one has a relevant fraction of the whole portfolio value.

- MtM_j: the *mark to market*, or present value, of the instruments issued by the *j-th* issuer, then $MtM_j = \sum_{i=1}^{I(j)} MtM_{i,j}$, where $I(j)$ is the number of instruments hold in the portfolio and issued by issuer j.

- DP_j: the default probability of the issuer *j-th*.

- R_j: the rating level of the issuer j-th at the evaluation time. Usually (it is merely a convention) the lower the rating code, lower the default probability. Usually the ratings are useful tools to group the default probability levels, hence one can write $DP(R_j)$ instead of DP_j to make explicit this mapping process.

- rr_j: the *recovery rate* for the *j-th* issuer once the default event happens. This value is a fraction in the range $[0, 1]$, but usually it is quite close to 35%, 40% for bond instruments, 0% for equity instruments. In the practice, we can not have an estimated recovery rate for each issuer, before it defaults, then the recovery rates are grouped by historical data and clustering all the default events by sector (financial, corporates, govies) and / or geography. We indicate with $s(j)$ the sector to which the issuer j belongs, we can write more explicitly $rr(s_j)$. The complement to 1 of the recovery rate is the loss given default (LGD in short), $lgd_j = 1 - rr_j$.

- L_j: the loss due to the default of the *j-th* issuer. Here we do not focus on the joint stochastic dynamics of market (interest rate, forex exchange, equity prices) and credit (default, spread) risk factors, hence we simply use the expression $L_j = MtM_j \cdot lgd_j \cdot F_t$, where F is the forward factor over a time horizon t and is equal to 1 at the default time is the MtM is supposed not to change significantly. Hence here L is not a random variable, but the loss once the default occurs. Furthermore, as usual from the *practitioner's point of view* as well as within the Basel regulation models, we replace the random loss given default fraction with its expected value LGD. Therefore, L can be considered as an expected loss, conditioned to the default event.

- RL: number of rating levels.

- $PL_j(rl)$ is the profit (or loss) for the holdings in the j-th issuer if its rating moves from the current level to the level $rl = 1, 2, \ldots, RL$. Practically, a further mapping process is performed, and for each combination (s, rl_1, rl_2) of rating migration for a given sector s a spread movement is established, let be $\Delta spread_{\{s, rl_1, rl_2\}}$. We stress the fact that, since the lgd_j do not differ significantly from a j to another, the vector of the losses L_j may be viewed as almost proportional to the vector of the mark to markets MtM_j.

With the above equipment, we can finally write the portfolio loss due to the default events, let $Loss_P$

$$Loss_P = \sum L_j \cdot \mathbb{1}_{D(j)},$$

being $\mathbb{1}_{D(j)}$ the *indicator* function of the default event. Despite the very compact expression, the calculation process of the risk figures (quantile, expected shortfall, etc.) are very involved, because of the complex parameters estimation process one has to set and to maintaining, and for the dependency structure between the issuers defaults. The *IDR* measure is defined as $VaR(Loss_P, 99.9\%, 1Y)$.

If we are interested to the broad profits and losses profile that we could observe in the portfolio because of the migration events we have

$$PL_P = \sum PL_j\,(R_j)\,.$$

In the above expression most of the complexity is in the chain of possible migration events, the associated PL, the dependency between the different issuers migration.

The above quantities are the key point of our application in the next section. At the end, the random variable PL_P is a discrete one. How many possible outcomes could it have?

Given the J issuers, and given the number RL of rating levels, we have that the cardinality of the outcomes of the random variable PL_P equals RL^J, i.e. $\#\{\text{outcomes}(PL_P)\} = RL^J$.

Here we are analyzing the issue by a strict *cardinality* perspective of the points in the space Ω where the elementary events ω_i take place. Of course we could have different events with the same numerical value of PL_P or some events that have very negligible probability. Anyway, with practical cases parameters it comes out a number of outcomes that can not be dealt satisfactorily in an exhaustive fashion.

Here *exhaustive* means that we could theoretically calculate the exact distribution of the random variable PL_p, i.e. its outcomes $\{x_i\}$ and the related probability masses $\{q_i\}$, and then to obtain the quantile by properly cutting the cumulative (discrete) distribution function at a given level.

With $J = 20$, $RL = 10$ we have 10^{20} possible outcomes. Obviously this cardinality cannot be actually managed due to 3 main reasons:

1. *computational*, i.e. to calculate all the $\{q_i\}$ probability masses;

2. *ordering*, i.e. to order the $PL_P(i)$ to get the quantile;

3. *storage*, i.e. to write in a database the whole input-output combinations for auditability purposes.

But thanks to the IDR new regularization the number of outcomes decreases dramatically, hence one can wonder if some exact calculation can be performed. Let us recall that in this case we have "only" 2^J elementary outcomes. With the above parameters $2^{20} \simeq 1,000,000$. Then at least theoretically we can try to face it by exploiting its feasibility, computational time and so on.

4 Evaluation approaches

In this section we will face our aim to determine the value of the minimum capital required to stand up to losses due to the 99.9% worst possible scenarios of issuers defaults with shares in the portfolio considered.

This part will be divided in two main subsections that correspond to the description of a numeric algorithm for a relatively small portfolio dimension and the description of a simulation algorithm for a portfolio with larger dimension. Because of the heuristic approach for the larger dimension portfolio, the second subsection will be integrated by a statistical comparison.

For both cases it will be considered a portfolio consisting of J issuers; J will be the dimension of the problem. Practical experience has shown that usually portfolios satisfy some characteristics on the loss vector $L = [L_1, L_2, \ldots, L_J]$ and

on the default probability vector $DP = [DP_1, DP_2, \ldots, DP_J]$, i.e. the vector which has as components the default probabilities of single issuers, that are summarized by the following properties:

- We have an usual *concentration effect* in the asset allocation (e.g. more in the govs bonds, less in the corporate bonds). Then we can heuristically assume that the loss determined by the default of the 10% of the issuers among the larger ones, should be at least the 90% of the maximum possible loss in the portfolio, $\sum_{j=1}^{J} L_j$. Here the values 10% and 90% are an example, and other similar assumptions on the concentration may be made. As we outlined in section 3, we have that $\mathbf{L} \approx c \cdot \mathbf{MtM}$, i.e. the 2 vectors have approximatively the same distribution. Hence from now on we simply use $\mathbf{L} = (L_j)$;

- L can be satisfactorily approximated by a *Beta distribution*, This approximation is justified since the global loss $L = \sum L_i$ is bounded, which is coherent with the Beta distribution characteristics and, moreover, by updating its parameters, the distribution behavior can be adapted to fit different loss shapes. It is also worth to mention that the Beta distribution is rather popular within financial applications to describe the loss given default quantity, which implies that it is a widely accepted random model in the credit risk management practice. We would also like to underline that the loss is the sum of weighted, by the portfolio fractions, loss given default. From this point of view, one can also consider different probability distribution alternatives, as the Gamma or Log-normal ones, nevertheless they are more suited for unbounded losses cases, as, e.g., when considering the operational risk field, as pointed out in [23].

- Issuers with greater L_j have a lower default probability DP_j, because of the prudent asset allocation of the bank's portfolio.

- Default probabilities are in the range $[0, 10\%]$.

It is worth to mention that such empirical facts admit some exceptions, nevertheless the above points have some rather intuitive rationale. Typically, if we look at the bond portfolio of a typical European commercial bank, it is characterized by long positions with a relevant part of government bonds, and, in this category, the first holding is for the national country bonds with *good tradition*, e.g., *BTP* for Italy, *bonos* for Spain, *bunds* for Germany, etc. Moreover, all the banks have some limits in the asset allocation policies, where some strict upper bounds are assigned for instruments and issuers with low rating, such as B, BB, etc. Furthermore, some limits are given also by sector, where corporate bonds, typically with higher default probability, cannot exceed a given threshold. Hence we have the double concentration effect mentioned above. As an example, we exhibit the portfolio composition, as of 2016, December, of one of the largest Italian banks. For the sake of clarity, here the weights are relative weights, and the total value of portfolio is about 20 billions €:

- Numbers of issuers in portfolio: about 90;

- Weight of the first bon issuer, or *Italian Republic* issuer: 84%;

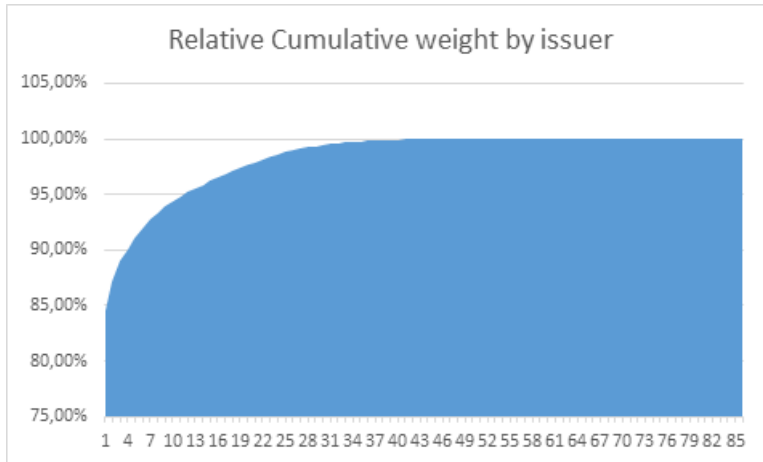- Weight of the first 5 holdings: 91%.

Figure 1: Plot of the cumulative weight of the bond issuers.

Figure 1 represents the plot of the cumulative weight of the bond issuers. Due to the huge variety of banks and issuers in the market, there is not a comprehensive empirical analysis in the literature, but several specific studies, most of which have been conducted by central banks and financial authorities that have access to data where generally there is not a granular disclosure, namely the bank portfolios holdings. A very recent deep analysis can be found in [21].

These portfolio management techniques explain both the concentration effect and the negative relationships between the PDs values and the exposures sizes. To this extent, let us recall that the market spread values for bonds can be misleading in the perception of the default probability, namely for the government bonds. In fact the FRTB regulation strictly prescribes to do not use the market, risk neutral, PDs implied by the spread, but instead the statistical real world estimation, see, in [7], the 352 BCBS paper, par.186.(s). Therefore, the high spread levels and volatilities characterizing some EU countries, where the banks have the most relevant holdings, do not imply high PDs in the DRC calculation, namely: the real probabilities are very low, hence matching the negative relationship principle.

To obtain a vector $\mathbf{L}$ satisfying these properties, for each single issuer $j \in \{1, 2, \ldots, J\}$ we simulate the values of L_j through a beta distribution $f(\alpha, \beta)$, with parameters $\alpha = 1/15$ and $\beta = 5$, of the default probability, DP_j, that will be proportional to its L_j.

In figure 2 are summarized the simulation steps for $\mathbf{L}$ and the default probabilities.

Instead, in figure 3 we split the range for the loss values in 40 sub-segments and represented the frequencies of occurrence of 1000 losses L_j, simulated by a $\text{Beta}\left(5, \frac{1}{15}\right)$, and we did the same for the default probabilities DP_j, which are inversely proportional to the value assumed by the correspondent L_j.

Figure 4 represents the scatter plot of $J = 200$ points of a single portfolio simulation for the couple (DP, L) to show the tendency of an inverse relation between the amount of losses and the default probabilities.
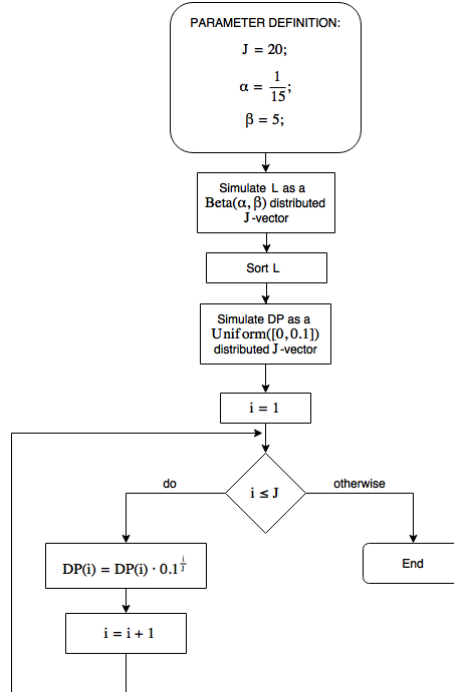
9

PARAMETER DEFINITION:

$J = 20;$

$\alpha = \dfrac{1}{15};$

$\beta = 5;$

Simulate **L** as a Beta(α, β) distributed J-vector

Sort L

Simulate DP as a Uniform($[0, 0.1]$) distributed J-vector

$i = 1$

$i \leq J$

do

otherwise

$DP(i) = DP(i) \cdot 0.1^{\frac{i}{J}}$

$i = i + 1$

End

Figure 2: Loss and Default Probability simulation.

Frequency histogram for losses simulated by a beta(1/15,5)
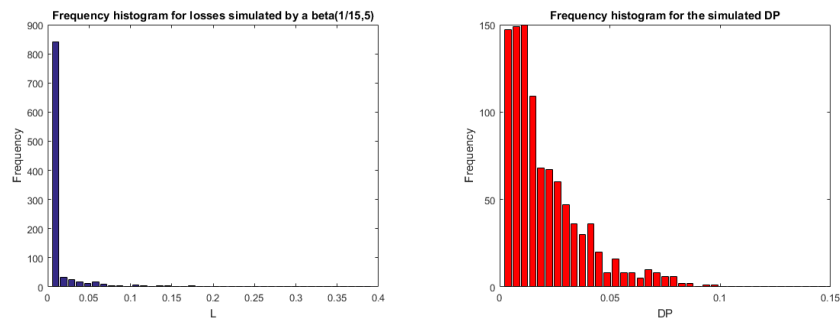
Frequency histogram for the simulated DP

Figure 3: Histogram representing the frequency of occurrences of 1000 values of L and DP. We remark the fact that banks have different portfolios from each other, and that the portfolios are time changing, therefore taking into consideration simulated portfolios is not less accurate than considering samplings that are relatively small w.r.t. the wide empiric variety.
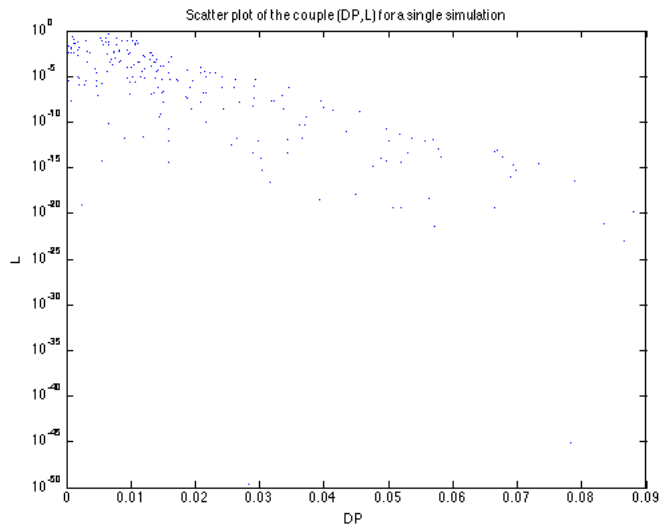
Figure 4: Scatter plot of a single simulation of $J = 200$ points of the couples (DP, L) with logarithmic scale for the y-axis.

4.1 Determination of the minimal capital required through comprehensive exact approach

Let us start from a simple case. We consider a portfolio P consisting in shares of $J = 20$ issuers, assuming that the loss vector L and the default probability vector DP are such that P satisfies the properties derived by practical experience.

Since every issuer is associated to a DP_j, we can compute the conditional default probabilities determined by defaults of combinations of issuers. In this paper we will assume that the issuers are uncorrelated, therefore it is the starting point for further analysis. The reason for this independence assumption on the issuers is mainly motivated by the fact that this study is meant to emphasize the calculation procedures and the machine execution effort, instead of the statistical model that binds the issuers.

Having $J = 20$ issuers in the considered portfolio, implies that the number of possible conditional default scenarios is

$$N = 2^J = 1048576,$$

and if we have n issuers that will default, the number of possible combinations is given by the binomial coefficient $\binom{J}{n}$. To each default scenario we can associate its conditional default probability CDP. For n issuers that will occur default, we consider the index set I with dimension n such that each element in I represents the index of a issuer that will face default in the simulated scenario. Then the indexes for the solvent issuers belong to the set

$$\{1, 2, 3, ..., J\} \setminus I.$$

11

The Conditional Default Probability associated to the index set I is

$$CDP(I) = \prod_{i \in I} DP(i) \cdot \prod_{\substack{1 \leq j \leq J \\ j \notin I}} \big(1 - DP(j)\big), \tag{1}$$

and its corresponding loss is

$$Loss_P(I) = \sum_{i \in I} L(i). \tag{2}$$

We want to determine the 99.9% worst case, so we sort all the possible $Loss_P$ in descending order as elements of the vector $\mathbf{Loss_P}^{\mathrm{Dec}}$, and the CDP with the same order as the sorted $Loss_P(I)$ as elements of the vector $\mathbf{CDP}^{Loss_P\text{-sort}}$. Therefore the first element of the vector for the reordered $\mathbf{Loss_P}^{\mathrm{Dec}}$ is the loss determined by the default of every issuer, i.e.

$$\mathbf{Loss_P}^{\mathrm{Dec}}(1) = Loss_P\big(\{1, 2, \ldots, J\}\big)$$
$$= \sum_{i=1}^{J} L(i),$$

and the first element of the $\mathbf{CDP}^{Loss_P\text{-sort}}$ reordered as the $\mathbf{Loss_P}^{\mathrm{Dec}}$ is

$$\mathbf{CDP}^{Loss_P\text{-sort}}(1) = \prod_{i \in 1}^{J} DP(i).$$

To provide an example, let us assume that the issuer with lower L is the second one, then the second element for the sorted $\mathbf{Loss_P}^{\mathrm{Dec}}$ would be the summation of all the issuers L except the one of the second issuer, and the second element for the reordered $\mathbf{CDP}^{Loss_P\text{-sort}}$ would be

$$\mathbf{CDP}^{Loss_P\text{-sort}}(2) = \prod_{\substack{1 \leq i \leq J \\ i \neq 2}}^{J} DP(i) \cdot \big(1 - DP(2)\big),$$

and so on for all the 2^J possible scenarios.

The IDR we are searching for is the element of $\mathbf{Loss_P}^{\mathrm{Dec}}$ corresponding to $\hat{k} - 1$, where $\hat{k}$ is the lower index such that the cumulative probability satisfies

$$\sum_{i=1}^{\hat{k}} \mathbf{CDP}^{Loss_P\text{-sort}}(i) > 0.1\%.$$

Since CDP is a discrete variable, we can not consider the usual definition of quantile of a continue random variable, instead we read the 99.9% quantile of the discrete variable CDP as the higher index $\hat{k}$ s.t.

$$\mathbb{P}(Loss \leq \mathbf{Loss_P}^{\mathrm{Dec}}(\hat{k} - 1)) \leq 99.9\%,$$

where $Loss : \Omega \to \mathbb{R}^+$ is a discrete random variable which takes values accordingly to (2) for all the possible combinations in $\Omega = \{0, 1\}^J$, where ones and zeros represent default and non-default for each issuer, e.g. an outcome

J	Time	normalized IDR
6	$0.073''$	58.47%
7	$0.038''$	69.97%
8	$0.068''$	69.11%
9	$0.141''$	61.76%
10	$0.267''$	43.02%
11	$0.465''$	97.23%
12	$0.910''$	70.96%
13	$1.633''$	76.83%
14	$3.188''$	80.32%
15	$6.22''$	63.85%
16	$12.21''$	82.01%
17	$26''$	84.21%
18	$52''$	59.4%
19	$1'\,45''$	41.58%
20	$3'\,32''$	59.79%
21	$7'\,4''$	51.56%
22	$14'\,2''$	96.31%
23	$27'\,48''$	80.47%
24	$55'\,31''$	32.60%
25	$1h\,41'\,41''$	33.64%

Table 1: Table representing the computation time and the normalized 99.9% IDR equal to IDR/$\sum_{i=1}^{J} L(i)$, later on it will be denoted as y, for portfolios simulated with different dimensions J.

$\omega \in \Omega = (0,0,0,1,0,0,1)$ represents the default of the fourth and seventh issuer in a $J = 7$ dimensional portfolio. Therefore for simplicity of notation we consider this approximation as the desired percentile, i.e.

$$\text{IDR}(99.9\%) = \textbf{Loss}_\textbf{P}^{\text{Dec}}(\hat{k} - 1). \tag{3}$$

In figure 5 we have summarized the *steps for the exact evaluation* of the IDR, i.e. to the $Loss_P$ corresponding to the 99.9% worst cases.

To conclude the analysis of the deterministic approach, let us underline that the exact algorithm we have presented, satisfactorily works with low dimension portfolios and uncorrelated issuers. While, if the dimension J representing the number of issuers starts to be greater than 30, then the computational efforts needed to evaluate the IDR for the 99.9% starts to be rather demanding, since the algorithm complexity grows exponentially in J, i.e. $T(J) = O(2^J)$. Table 1 relates the dimension of the simulated portfolios to the amount of time taken by the algorithm to run using a CPU based on a 2,5 GHz Intel Core i5 unit.

4.2 Determination of the minimal capital required through heuristic technique HR

As we have seen in the previous subsection, increasing the portfolio dimension the execution time for the exact algorithm becomes prohibitive. For this reason, for an higher dimensions we simplify the complexity of the IDR evaluation problem through an estimate from a regression model. Let us denote as J^*
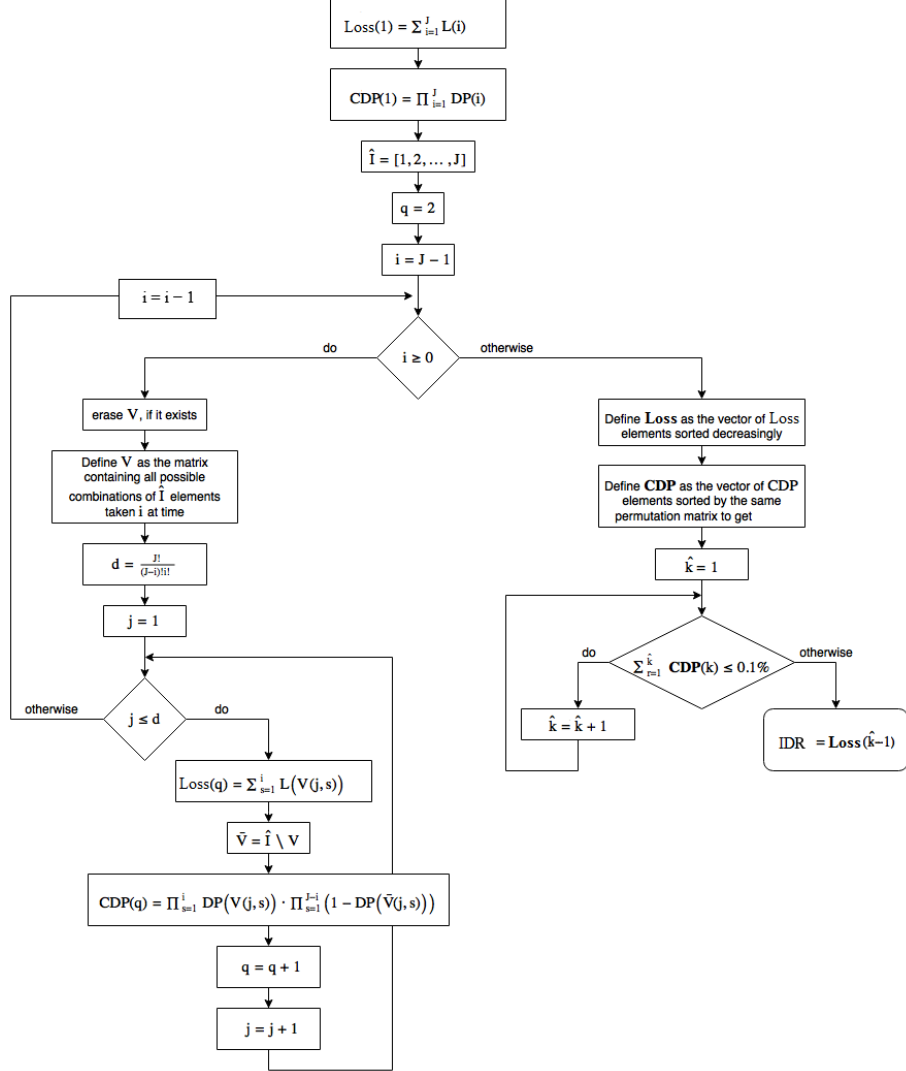
Figure 5: Evaluation procedure for the IDR, that corresponds the $Loss_P$ in to the 99.9% worst cases. For the sake of clarity, let us divide the graph in three main parts: before the $i \geq 0$-*cycle*, *during* and *after* the cycle itself:

1. The initialization of the variables *Loss* and *CDP* as vectors with $\sum_{i \in \{0,1,...,J\}} \binom{J}{i} = 2^J$ elements. Moreover the definition of some counting variables for the next cycles.

2. Repeat for $i = J-1, \ldots, 1, 0$, then the cycle consists in defining the matrix V whose rows are the possible combination of i banks failing. Then define all the elements in the vector *Loss* as all the possible losses corresponding to the combination of failing banks. For each loss, we determine the probability to face this amount of default defining them as the elements of the vector *CDP*.

3. Search for the desired percentile and define it as *IDR*.

the higher dimensions in order to mark the difference between this estimation problem and the previous one. We also denote as P^* the portfolio consisting in shares of J^* issuers, and will be referred as *Portfolio Target*.

Considering n simulated J dimensional portfolios, we estimate the relationship between y, ratio of IDR(99.9%) with respect to the maximum possible $Loss_P$ given by default scenarios, i.e.

$$y = \frac{\text{IDR}(99.9\%)}{\sum_{i=1}^{J} L(i)}, \tag{4}$$

and two concentration indexes. We remark the fact that y is a relative IDR which is a dimensionless index. Indeed, the basic idea behind the evaluation of y is to start estimating the IDR of the Portfolio Target P^* through the concentration indexes and then switching to a dimensionless scale w.r.t. its dimension J^*. To be more precise, supposing that the quantile is a well estimable fraction, given by the concentration curves of L and DP, we extrapolate the behavior observed in the lower dimension $J < J^*$ to adapt it to the Target problem in dimension J^*. We consider the concentration indexes that measure how the total $Loss_P$ and the sum of all the default probabilities are divided through the issuers. Let us explain how we get these concentration indexes. First of all we sort the issuers by their losses L. Then, since to each issuer is associated a L and a DP, from the reordering we obtain two vectors:

1. $\mathbf{L}^{Inc}$, which has as elements the L arranged increasingly,

2. $\mathbf{DP}^{L\text{-sort}}$, which has as elements the DP corresponding to the losses L with same index in the vector $\mathbf{L}^{Inc}$.

Let us denote by $Q_1^{(p)}$ and $Q_2^{(p)}$ the concentration indexes for $\mathbf{L}^{Inc}$ and $\mathbf{DP}^{L\text{-sort}}$ given by the formulas

$$Q_1^{(p)} = \frac{\sum_{i=1}^{\lceil p \cdot J \rceil} \mathbf{L}^{Inc}(i)}{\sum_{i=1}^{J} \mathbf{L}^{Inc}(i)}, \quad Q_2^{(p)} = \frac{\sum_{i=1}^{\lceil p \cdot J \rceil} \mathbf{DP}^{L\text{-sort}}(i)}{\sum_{i=1}^{J} \mathbf{DP}^{L\text{-sort}}(i)}, \tag{5}$$

where p is the percentage corresponding to the portion of issuers which will cause a lower loss in the portfolio w.r.t. the other $(1-p)J$ issuers. We remark the fact that the index $Q_1^{(p)}$ represents the ratio of minimum possible loss in the portfolio $\sum_{i=1}^{J} L_j$ that would occur in the case that the portion p of issuers which would cause the lower loss in the portfolio will face the default event. Respectively, $Q_2^{(p)}$ is the concentration ratio of the default probabilities of the issuers which would determine lower losses in the portfolio. Table 2 shows the two concentration indexes Q_1 and Q_2 corresponding to a simulated portfolio for every percentage p.

In the regression we consider as regressors the concentration indexes $Q_1^{(90\%)}$ $Q_2^{(75\%)}$, i.e. through linear regression the model would be

$$y = \beta_0 + \beta_1 Q_1^{(90\%)} + \beta_2 Q_2^{(75\%)} + \epsilon, \tag{6}$$

where the slopes β_1 and β_2 and the intercept β_0 are the unknown parameters, the normalized quantile y is the dependent variable, the rate index, Q_1 and Q_2 are the independent variables and ϵ represents the estimation error. Namely, the

p	$\lceil p \cdot J \rceil$	$Q_1^{(p)}$	$Q_2^{(p)}$
$(0\%, 5\%]$	1	0.0000%	11.64%
$(5\%, 10\%]$	2	0.0000%	22.60%
$(10\%, 15\%]$	3	0.0000%	33.39%
$(15\%, 20\%]$	4	0.0001%	45.12%
$(20\%, 25\%]$	5	0.0002%	56.65%
$(25\%, 30\%]$	6	0.0003%	64.70%
$(30\%, 35\%]$	7	0.0005%	70.12%
$(35\%, 40\%]$	8	0.0006%	77.83%
$(40\%, 45\%]$	9	0.0008%	82.13%
$(45\%, 50\%]$	10	0.0023%	82.24%
$(50\%, 55\%]$	11	0.0076%	82.95%
$(55\%, 60\%]$	12	0.0438%	87.48%
$(60\%, 65\%]$	13	0.18%	89.74%
$(65\%, 70\%]$	14	1.16%	93.26%
$(70\%, 75\%]$	15	3.51%	94.04%
$(75\%, 80\%]$	16	7.79%	95.86%
$(80\%, 85\%]$	17	13.60%	97.72%
$(85\%, 90\%]$	18	25.85%	97.80%
$(90\%, 95\%]$	19	55.24%	99.24%
$(95\%, 100\%]$	20	100%	100%

Table 2: Table representing $Q_1^{(p)}$ and $Q_2^{(p)}$ values corresponding to a simulated $J = 20$ dimensional portfolio, for every p ratio.

underlying idea is that a general portfolio satisfying the characteristics arising from practical experience can adequately be represented by these two concentration indexes; and therefore, for high-dimension portfolios, it suffices to make regression on $Q_1^{(p_1)}$ and $Q_2^{(p_2)}$, for proper p_1 and p_2.

The choice of such a percentage p for $Q_1^{(p)}$ is embedded in the hypothesis of concentration above the construction of **L**. To select such p we simulated several curves $Q_1 : [0, 1] \to [0, 1]$, $p \mapsto Q_1^{(p)}$ (as we did in figure 6) and we considered the tangent lines $T_{\bar{p}}$ to the graph described by Q_1 in different points $\bar{p} \in (0, 1)$

$$T_{\bar{p}}(x) = \frac{\mathrm{d}Q_1^{(p)}}{\mathrm{d}p}(x - \bar{p}) + \bar{p}. \tag{7}$$

In particular we considered the slopes of the tangent lines and it turns out that for $p \in (0, 0.85]$ these slopes were close to zero and too much less than 1, therefore the variations $\frac{\Delta Q_1^{(p)}}{\Delta p}$ were exiguous, which means that choosing $\bar{p} = 0.1$ or $\bar{p} = 0.7$ would result in $Q_1^{(\bar{p})}$ very close to each other; on the contrary for $\bar{p} \in [0.95, 1)$ the slopes are too conspicuous and therefore different simulations could give values of $Q_1^{(\bar{p})}$ too distant. For these reasons we choose a middle value $p = 0.9$, so that $Q_1^{(p)}$ would have been appropriate and meaningful for the regression. We followed the same theory to choose the percentage p for $Q_2^{(p)}$. This time the only care needed was p not being close to 1 since the issuers with higher L have very poor DP, so the variations $\frac{\Delta Q_2^{(p)}}{\Delta p}$ would have been close to
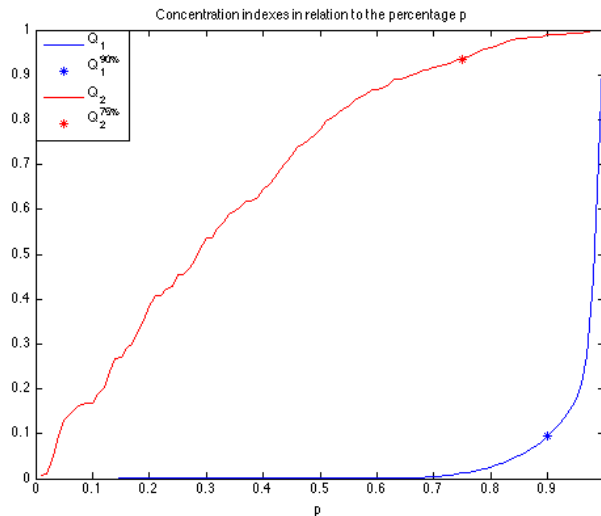
Figure 6: Plot of the concentration indexes of a simulated portfolio.

zero and the values of the regressors wouldn't have been very significant for the regression.

In figure 6 are represented the concentration indexes for a portfolio simulated as explained in figure 2; the chosen $Q_1^{90\%}$ and $Q_2^{75\%}$ are marked.

A linear regression model could be inappropriate to estimate the rate y. In fact, it would not consider the upper and lower boundaries, hence, to avoid predictions outside the interval $[0, 1]$, we apply the *Logistic Quantile Regression*(LQR). There are several application of the LQR to systemic risk analysis, see [26] for a study concerning the forecast of recovery rates. Moreover, LQR is widely used in many statistical fields, as, e.g., in Economics, as well as in Ecology, Meteorology, Biomedical sciences, etc., see, e.g., [11, 18], and references therein.

Hence through the logistic formula for the dependent variable y

$$\text{logit}(y) = \log\left(\frac{y}{1-y}\right)$$

we bound the estimated values between 0 and 1. Inverting the logit function we get the regression model as

$$y \sim \frac{1}{1 + \exp\left(-\left(\beta_0 + \beta_1 \, Q_1^{(90\%)} + \beta_2 \, Q_2^{(75\%)} + \epsilon\right)\right)}. \tag{8}$$

We start considering the simulation of $n = 50$ portfolios composed by shares of $J = 12$ different issuers.At this first introduction level n does not play a specific role, we just use $n = 50$ as it is enough for an estimation of the 2 parameters in the regression. Each of them is simulated by the procedure explained in the previous subsection. So we compute n couples of concentration indexes $Q_1^{(90\%)}$ and $Q_2,^{(75\%)}$, and rate indexes y, and through the logit regression model we estimate the parameters β_0, β_1 and β_2. In figure 7 are plotted the triplets
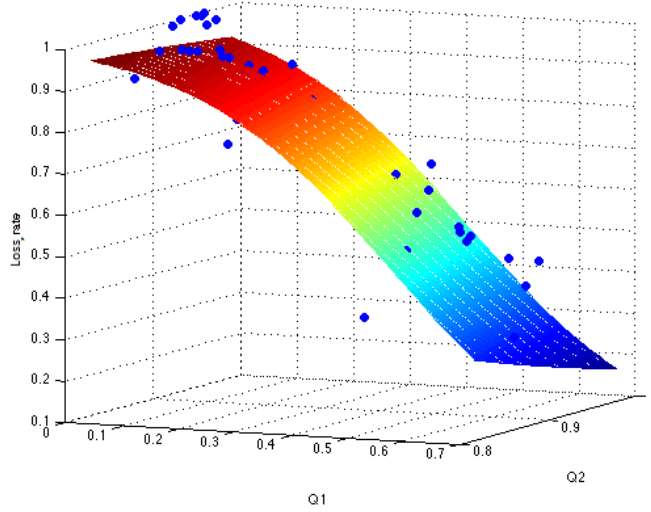
17

Figure 7: Plot of the logistic regression model and of the data obtained by the simulation of $n = 50$ portfolios.

$(Q_1^{(90\%)}, Q_2^{(75\%)}, y)$ corresponding to the n portfolios, and the regression model.

Once the parameters β_0, β_1 and β_2 are estimated, they can be used to estimate the value of the rate y^* corresponding to the 99.9% worst case for the portfolio P^* with dimension J^* as

$$y^* = \text{logit}^{-1}\left(-\beta_0 - \beta_1 Q_1^{(90\%)^*} - \beta_2 Q_2^{(75\%)^*}\right), \qquad (9)$$

where $Q_1^{(90\%)^*}$ and $Q_2^{(75\%)^*}$ are the concentration indexes for P^*. Consequently the estimated Loss for P^* is

$$\text{IDR}^*(99.9\%) = y^* \cdot \sum_{i=1}^{J^*} L^*(i), \qquad (10)$$

where $L^*(i)$ is the loss determined by the default of the i$^{\text{th}}$ issuer in the portfolio P^*.

In figure 8 through a flow chart are summarized the steps for the HR model and the estimation of the IDR for P^* in the 99.9% worst cases.

Monte Carlo simulation and IDR estimation Let us consider a portfolio P with J independent issuers. As before, a determined loss value $L(j)$ and a default probability $DP(j)$ are assigned to each issuer $j = 1, \ldots, J$, see figure 2 for the detailed simulation procedure. The Monte Carlo (MC) method consists in the simulation of N default scenarios which can be summarized by the following steps:

1. Start simulating a uniform-randomly based matrix $\mathbf{U}$, with values in $[0,1]^{J \times N}$, namely each of its elements $U(j,n)$, for $j \in \{1, \ldots, J\}$ and $n \in \{1, \ldots, N\}$, corresponds to the j-th issuer in the n-th scenario;
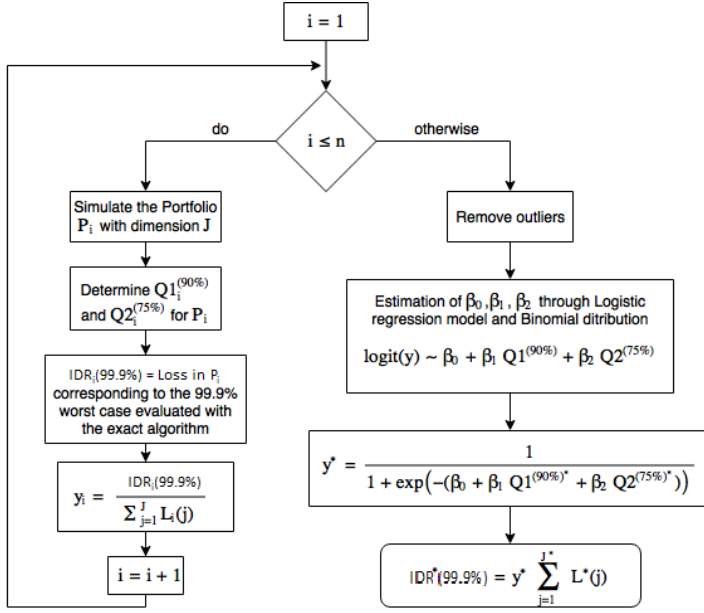
18

Figure 8: Flow chart of the steps for the estimation of the Portfolio Target IDR in the 99.9% worst cases. Given n, number of regression Portfolio that have to be simulated, J and J^* dimension for the regression Portfolios and the Portfolio Target respectively, and the $L^*(j)$ for every issuer of the Portfolio Target and the $DP^*(j)$, with $j = 1, 2, \ldots, J^*$.

2. Then, for each scenario $n = 1 \ldots, N$, we have the j-th issuer default as soon as $U(j, n) < DP(j)$;

3. The loss in the portfolio P, for the n-th simulated scenario, is given by the sum of the $L(j)$ s.t. $U(j, n) < DP(j)$, i.e.

$$Loss^{MC}(n) = \sum_{j=1}^{J} L(j)\, \mathbb{1}_{\{U(j,n)<DP(j)\}}, \quad \text{for } n = 1, \ldots, N. \tag{11}$$

We use the empirical quantile as the MC estimate of the theoretic quantile, i.e. the estimated $\text{IDR}^{MC}(99.9\%)$ is the $(N/1000)$-th greatest $Loss^{MC}$.

By construction, we have that the random variable $Loss$ is s.t.

$$\mathbb{P}\big(Loss \leq \text{IDR}(99.9\%)\big) \approx 99.9\%, \tag{12}$$

where $\text{IDR}(99.9\%)$ is the 99.9% percentile that we aimed to estimate. Therefore the sample $p = 99.9\%$ quantile $\text{IDR}^{MC}(p)$ satisfies the following equation

$$\hat{\sigma}_N^2 = Var(\text{IDR}^{MC}) \approx \frac{p\,(1-p)}{(N+2)\,f^2(\text{IDR})} + O(1/N^2), \tag{13}$$

with f probability density function (PDF) of $Loss$, see, e.g., [12]. In figure 9 is plotted f, the PDF estimated from the simulated data $Loss_1^{MC}, \ldots, Loss_N^{MC}$, with $N =$1,000,000. The estimation is based on a normal kernel function. We estimate the confidence interval with 99.9% confidence level as

$$CI_N^{MC}(99.9\%) = \text{IDR}^{MC} \pm 3.0902\,\hat{\sigma}_N. \tag{14}$$
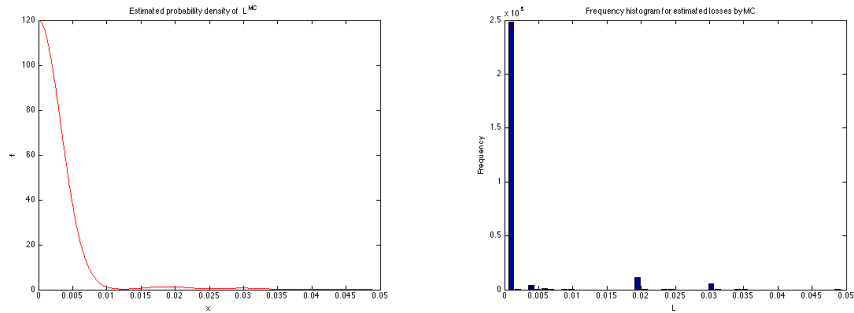
Figure 9: Estimated PDF of *Loss* from the MC simulated data represented in the adjacent frequency histogram.

Comparison between HR Model and Exact Algorithm We consider $J^* = 18$ as the dimension of the portfolio target. Our aim is to compare the HR model with the exact algorithm and the MC model and see how far is the prediction for the IDR obtained by the HR from the correct IDR obtained by the exact algorithm. So we simulate $n = 32$ regression portfolios with dimension $J = 12$ (at the end of this paragraph our choice for n and J will be explained).

The portfolios are simulated as shown in figure 2, namely they share the characteristics of beta distributed losses and default probabilities are inversely related to the correspondent losses. The choices of J^*, n and J are arbitrary, with the premise that $J < J^*$; but, moreover we have to explicit that n and J condition the time spent to realize the regression, since the regression model has complexity $T(J) = n \times O(2^J \ln(2^J)) \approx n \times O(2^J)$. From the logistic regression we get the parameters

$$\beta_0 = 3.2945,$$
$$\beta_1 = -5.6408,$$
$$\beta_2 = -0.8441.$$

and therefore the shape of the surface for the triplets $(Q_1^{(90\%)}, Q_2^{(75\%)}, y)$ would be similar to the one represented in figure 7. Afterwards, these parameters are used to estimate the IDR of $N = 10$ portfolios target in order to compare them with the IDR evaluated by the exact algorithm.

Alongside this we simulate $N^* = 10,000$ and $N^{**} = 100,000$ default scenarios for the MC method and we estimate the IDR and the CI, as shown in the previous paragraph.

In figure 10 are plotted the rate values y^* obtained by the exact algorithm and the HR model, both corresponding to the N Portfolios Target, and in table 3 are shown the obtained values compared with the CI resulting from the MC simulations. The order of magnitude concerning execution time and accuracy of the iterations of the HR and the MC model are not clearly comparable, since HR is much faster, while MC is usually more accurate, especially for a great number of MC simulations N^{MC}. As to provide a meaningful comparison, we have simulated 20 portfolios with dimension $J^* = 18$. Figure 11 represents the execution times in a logarithmic scale and the estimation errors that we got by the HR method and the MC simulations for $N^{MC} = 10^3, 10^4, 10^5$, although
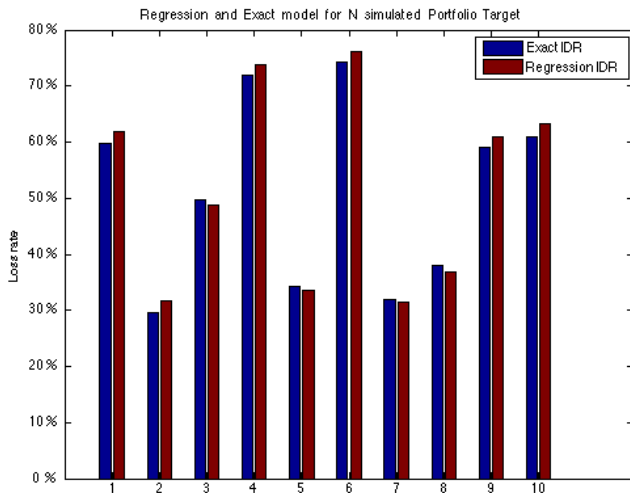
Figure 10: Histogram of the comparison between the values obtained with the exact algorithm and the HR model.

MC for $N^{MC} = 1000$ is not very significative, but we considered it in order to compare it with HR and have a similar order of magnitude for the error. We point out the fact that even though the real distribution of the IDR estimated by MC is not known, we build the order statistics in order to estimate the probability density relying on a normal kernel function. Then the estimate of the pdf was exploited to estimate the variance of IDR^{MC}, and enable us to construct the CI and compare them with the errors of the HR method.

For the purpose of comparing the expenditure in time we remark the fact that to compute the estimation for the parameters β_0, β_1 and β_2 it took 28 seconds, and less than a millisecond for the prediction of the IDR for each portfolio target, i.e. by applying the HR model. Instead, the evaluation of the IDR took approximately a minute. The computation time for the simulation with MC depends on the dimension J^* and on the number of simulations N^{MC}; the times listed in table 4 have to be compared with the millesimal time required by the HR model.

Moreover, increasing the dimension of the problem, the time to estimate the IDR with the HR approach remains almost the same, indeed previously we had seen that the computational complexity is $T(J^*) \approx n \times O(2^J)$, i.e. it depends on the chosen regression parameters n and J, instead of J^*. On the other hand, the time to evaluate the IDR with the exact approach of section 4.1 increases exponentially. This is due to the fact that the computational complexity for the exact algorithm is related to the input parameter J^*, since the mean complexity of the best possible sort of a k-dimensional vector is $O(k \ln(k))$, see [19, sec. 5.4, pp. 248–379]. Therefore, the resulting complexity is

$$T(J^*) = O(2^{J^*} \ln(2^{J^*})) = O(2^{J^*} J^* \ln(2)) \approx O(2^{J^*}). \qquad (15)$$

We have this order of computational complexity since we are using an exact algorithm that sorts the vector of the 2^{J^*} possible loss outcomes, then it computes
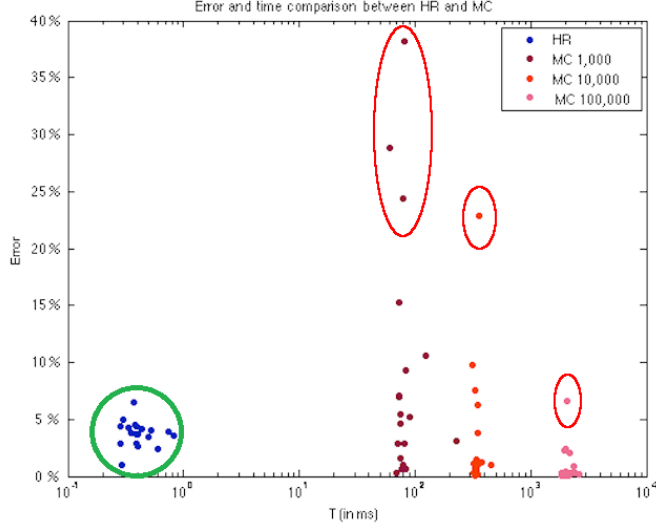
Figure 11: Comparison between HR and MC model that puts in relation the execution times T and the estimation errors E. The HR error is given by the absolute value of the difference of the exact IDR and the estimated IDR, while the MC error is equal to the width of the estimated confidence interval, i.e. equal to $6.1804\,\hat{\sigma}_N$. We marked with a red circle the outliers of MC, and stress the fact that even for a MC with $N^{MC} = 10^4$, which is almost equal to $2^{J^*} = 262144$, there are chances of outliers.

Ex. IDR	HR IDR	Delta	$\hat{\sigma}_{N^*}^{MC}$	$CI_{N^*}^{MC}(99.9\%)$	$\hat{\sigma}_{N^{**}}^{MC}$	$CI_{N^{**}}^{MC}(99.9\%)$
59.83%	61.89%	-2.06%	0.24%	[59.21%, 60.69%]	0.05%	[59.73%, 60.03%]
29.60%	31.58%	-1.97%	0.87%	[28.90%, 34.28%]	0.20%	[29.53%, 30.76%]
49.76%	48.77%	0.99%	0.23%	[49.20%, 50.63%]	0.06%	[49.51%, 49.89%]
71.85%	73.77%	-1.92%	0.15%	[71.59%, 72.52%]	0.03%	[71.76%, 71.95%]
34.23%	33.52%	0.71%	0.09%	[33.71%, 34.26%]	0.01%	[34.21%, 34.28%]
74.29%	76.09%	-1.79%	0.50%	[73.45%, 76.54%]	0.12%	[74.02%, 74.76%]
31.93%	31.47%	0.46%	0.01%	[31.90%, 31.97%]	0.004%	[31.93%, 31.95%]
37.99%	36.87%	1.11%	0.54%	[36.12%, 39.46%]	0.18%	[37.24%, 38.35%]
59.04%	60.97%	-1.92%	0.28%	[58.62%, 60.11%]	0.07%	[58.87%, 59.30%]
60.82%	63.36%	-2.53%	0.02%	[60.78%, 60.91%]	0.003%	[60.82%, 60.83%]

Table 3: Table containing IDR got from the HR model and the exact algorithm for a $J^* = 18$ dimensional portfolio, where Delta = HR IDR − Exact IDR. The $\hat{\sigma}_{N^*}$ and $\hat{\sigma}_{N^{**}}$ are estimated by equation (13) with $N^* = 10,000$ and $N^{**} = 100,000$, the CI are centered in the IDR values obtained by the MC estimation.

J^*	MC time	HR time	$\mathrm{MC_{time}/HR_{time}}$
15	3.83s	0.002s	1900
20	4.96s	0.004s	1250
25	5.86s	0.004s	1450
30	6.43s	0.003s	2150
40	8.64s	0.007s	1250

Table 4: Table containing the times to evaluate the IDR with the MC and HR methods and for different portfolio dimensions J^* (the parameters $\beta_0, \beta_1, \beta_2$ were estimated once and for all). In the last column we indicate how many times the HR method results faster than the MC method.

the CDP until it finally stops when it reaches the desired percentile.

An alternative approach for the exact algorithm is represented by the scanning of the tree of cases, i.e. an algorithm that firstly computes the greatest loss L_1, and its CDP, then the second greatest loss L_2 and its CDP, the third and so on, until the percentile 0.1% is reached. But for the choice of structuring the portfolios as shown in fig.2 it is necessary to consider circa the half of all possible combinations, and therefore it takes even more time than the previous exact algorithm.

Let us underline that we have decided to consider $J = 12$ and $n = 32$, w.r.t. $J^* = 18$, mainly because of the computational effort required by the HR and exact method. In particular, to make the HR more efficient w.r.t. the exact calculation, i.e. with smaller computational complexity

$$O(\mathrm{HR}) < O(\mathrm{exact}),$$

we need $n \times O(2^J)) < O(2^{J^*})$, hence

$$J^* - \log_2(n) > J,$$

see equation (15) to what concerns the complexity of the exact algorithm. Therefore, since $n = 32$ is a sufficient sample size, for $J^* = 18$, J has to be at least less than 14. In table 5 are shown the times spent to compute the regression model for various J and n, these times have to be compared with the time spent to execute the exact algorithm (in this case, with $J^* = 18$, it was spent approximately one minute, see table 1 for further cases).

The computational workflow. A summary Let us go back to fig.8 for a global review of the algorithm. In the application perspective, the accuracy and the execution time are very important, but we also need to have a clear description of the algorithm, its set-up complexity, the maintenance effort and so on.

Generally the "long" bank bond portfolios change quite slowly, hence the concentration and correlation measures between the exposures are very smoothly over time. Hence, we can split the workflow in 3 different levels. A first level is the general (near static) set-up, i.e. the definition of the general parameters, such as the number of simulation N, the regression technique, the size J for the benchmark portfolio, the α levels for Q_1 and Q_2, etc.

	n	
J	32	64
12	29"	53'
13	54"	1'56"
14	1'48"	3'37"
15	3'32"	7'17"
16	7'12"	13'50"

Table 5: The time to compute the regression does not depend on the Portfolio target P^*, and in particular not on J^*. This table reports the times spent to compute the regression for different J and n. We remark the fact that the regression time is a set-up which can be estimated once a month, and therefore since the meaningful time for the HR method is the one reported in the third column of table 4, for high Portfolio dimensions, the time saved is considerable.

The second level, with a periodic update, is given by the core parameters estimation, i.e. the IDR calculation for benchmark portfolios and the parameters (β_1, β_2) estimation by the regression procedures.

Finally the third level is the execution task, in other words the calculation of the estimate IDR, given the current actual portfolio parameters and the estimate β's.

As concerns the frequency of the above processes, the general set-up is usually triggered by any top management (risk control, risk committee) new guidelines or yearly. The periodic update could be run monthly or quarterly according to the portfolio dynamics. Finally the execution process is related to the bank regulatory constraints. If the bank has validated the internal model for the specific market risk, e.g. following a specific solution as the one underlined in [10] a daily calculation is usually required. Otherwise the bank can set its own reporting frequency.

To summarize, the proposal can fit quite easily the internal usual bank procedures and processes.

5 Extension to the correlated case and possible approaches

Until now only the case with uncorrelated issuers was considered. In the previous sections of this paper the outside-diagonal elements of the correlation matrix were set equal zero. Although this setting is useful to be the starting point to deal with the heuristic approaches to the IDR estimation problem, from a financial point of view this is unrealistic.

In the non-independent portfolio case, i.e. removing the issuers' independence assumption, also the elements outside the diagonal are non-zero. Correlations are not observable from the market, and depend on the underlying assumptions of the assets values, such as the Black and Scholes model, from which they can be obtained. However from a theoretical point of view the correlations are estimated from the CDS spreads of the the issuers: but usually this is impracticable also for the shortage of data. So often they are estimated from the equity prices. Furthermore computing the IDR for dependent portfolios is

more complex, and in particular unfeasible even for portfolios with not very huge dimensions due to the great execution time.

Therefore our aim is to build an adequate structure to apply an HR model. The technical problem underlying the new HR model is that we would not have to estimate only the coefficients β_0, β_1, β_2 of the previous (uncorrelated) model, but also all the coefficients concerning the correlations between the issuers. Since this is impracticable, we have to find less variables that properly represent the correlations between the issuers.

Before explaining our proposal, let us summarize what the FRTB regulations claims for the IDR model parameters calibration. The general point is that the high confidence level for the measure, i.e. 99.9%, along with the rather long time horizon, i.e. 1 year, does not allow for an easy IDR-back test procedure for both the default events of the financial instruments issuers and for the related correlations. Practically, this is no feasible, hence in the FRTB new framework some flexibility is allowed, see [7]. In particular, we have the following main points to be considered

- Correlations should be estimated over a 10 years periods, covering a stress period, hence allowing to consider increases of correlations during the financial crisis

- Correlations must be inferred from spreads or equity prices, with some proxies/benchmark where not available

- Banks must have clear policies and procedures that describe the correlation calibration process

- Previous correlations must be based on objective data

- A bank must validate its modeling approach for such correlations, namely it has to show that the provided *internal* method is appropriate for the detained portfolio, also with respect to the chosen systematic risk factors and associated weights

We would also like to point out the remark written by BCBS about the DRC validation, namely *Accordingly, the validation of a DRC model has to rely more heavily on indirect methods including, but not limited to, the stress tests, sensitivity analysis and scenario analysis, hence to assess both its qualitative and quantitative efficiency and robustness, particularly with regard to the model's treatment of concentrations.*

To be more explicit, banks must have sound procedures and processes to assess, build and update their models with proper documentation, even if, due to the very challenging goal of the IDR estimation, some flexibility is permitted.

We note that usually portfolios are comprised of issuers that are correlated each other in a similar way (this feature is due to the geographic consistence of portfolios). Therefore the outside-diagonal correlation matrix elements don't vary very much and we can consider the mean value of them as one of the regressors of the new HR model, let us call it $\widetilde{\sigma}$. Moreover in the regression we add the regressor D, corresponding to the index of dispersion of the outside diagonal elements. So, for each portfolio, simultaneously to the computation of $Q_1^{(90\%)}$ and $Q_2^{(75\%)}$, obtained as concentration indexes of L and DP respectively, we need to compute $\widetilde{\sigma}$ and D.

Therefore the new linear regression will be

$$y = \beta_0 + \beta_1 \, Q_1^{(90\%)} + \beta_2 \, Q_2^{(75\%)} + \beta_3 \, \widetilde{\sigma} + \beta_4 \, D + \epsilon, \tag{16}$$

where the dependent variable y is the normalized quantile, the intercept β_0 and the slopes β_i for $i = 1, 2, 3, 4$ are the unknown parameters, the rate index Q_1 and Q_2, the mean correlation $\widetilde{\sigma}$ and the dispersion index D are the independent variables, and ϵ represents the estimation error.

References

[1] Acerbi C., and D. Tasche (2002). On the coherence of the Expected Shortfall. *J. Bank. Financ. 26* (7), 1487-1503.

[2] Acerbi C., and B. Szekely (2014). Backtesting Expected Shortfall. *MSCI Research paper.*

[3] Basel Committee on Banking Supervision (1996). Amendment to the Capital Accord to Incorporate Market Risks, *Paper 119.*

[4] Basel Committee on Banking Supervision (2006). International Convergence of Capital Measurement and Capital Standards. *128.*

[5] Basel Committee on Banking Supervision (2009). Guidelines for computing capital for incremental risk in the trading book. *Paper 159.*

[6] Basel Committee on Banking Supervision (2009). Revisions to the Basel II market risk framework - final version. *Paper 158.*

[7] Basel Committee on Banking Supervision (2016). Minimum capital requirements for market risk. *Paper 352.*

[8] Benazzoli, C. and Di Persio, L. (2016). Default contagion in financial networks *Int. Journ. of Math. and Comp. in Simulation* Vol. *10*, pp.112-117

[9] Bonollo M. (2015). The Fundamental Review of Trading Book (FRTB): Revolution or (R)-Evolution ?. Available at *www.finriskalert.it.*

[10] Bonollo M., Di Persio, L., Mammi, L. and Oliva, I. (2017) Estimating the Counterparty Risk Exposure by Using the Brownian Motion Local Time *Int. Journ. of Appl. Math.and Comp. Science, 27* (2), pp. 435-447.

[11] Bottai M., B. Cai, and R.E. McKeown (2010). Logistic quantile regression for bounded outcomes. Statistics in Medicine. *Stat. Med. 29* (2), pp. 309-317.

[12] Chen E. J. (2002). Two-phase quantile estimation. *IEEE Xplore 1*, pp. 447-455. DOI: 10.1109/WSC.2002.1172916.

[13] Wilde T. (1997). Credit Risk+: A Credit Risk Management Framework. *CSFP* (Technical paper).

[14] David H. A., and H. N. Nagaraja (2005). *Order Statistics.* Wiley Ser. Probab. Stat.

[15] European Parliament (2013). *REGULATION (EU) No 575/2013 on prudential requirements for credit institutions and investment firms and amending Regulation.*

[16] Glasserman, P. (2012). *Risk horizon and rebalancing horizon in portfolio risk measurement.* Math. Finance, 22(2), pp. 215–249.

[17] Gordy M.B. (2003). A Risk-Factor Model Foundation for Ratings-Based Bank Capital Rules. *J. Financial Intermediation, 12.*

[18] Koenker R., and G. Bassett (1978). Regression quantiles. *Econometrica,*
46(1), pp. 33–50.

[19] Knuth D. (1988), *THE ART OF COMPUTING PROGRAMMING*: Volume 3. Sorting and Searching, Second Edition. Addison-Wesley. ISBN 0-201-89685-0.

[20] Laurenta J.P., M. Sestiera, and S. Thomasb (2016). Trading book and credit risk: how fundamental is the Basel review. *J. Bank. Financ., 73*(1).

[21] Manna M., and S. Nobili (2018). Banks' holdings of and trading in government bonds, *Bank of Italy Working paper, n.1166.*

[22] Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *J. Finance, 29(2)*, pp. 449-470.

[23] Moscadelli M. (2004). The modelling of operational risk: experience with the analysis of the data collected by the Basel Committee. *Bank of Italy working paper*, n. *517.*

[24] MSCI RiskMetrics Group (2007). *CreditMetrics technical document.*

[25] Predescu M. and S. Wilkens (2017). Default risk charge: modeling framework for the "Basel" risk measure. *J. Risk*, 19(4).

[26] Siao J.H., R.C. Hwang, and C.K. Chu (2016). Predicting recovery rates using logistic quantile regression with bounded outcomes. *Quant. Finance,* 16(5) pp. 777-792.

[27] Vasicek, O. (2002). The Distribution of Loan Portfolio Value. *Risk*, December.