The Climate Extended Risk Model (CERM)

Josselin Garnier *

December 17, 2020

This paper is directed to the financial community and focuses on the financial risks associated with climate change. It, specifically, addresses the estimate of climate risk embedded within a bank loan portfolio. During the 21st century, man-made carbon dioxide emissions in the atmosphere will raise global temperatures, resulting in severe and unpredictable physical damage across the globe. Another uncertainty associated with climate, known as the energy transition risk, comes from the unpredictable pace of political and legal actions to limit its impact. The Climate Extended Risk Model (CERM) adapts well known credit risk models. It proposes a method to calculate incremental credit losses on a loan portfolio that are rooted into physical and transition risks.
The document provides detailed description of the model hypothesis and steps. It is structured as follows: Part 1 is an introduction to the goals and main concepts. Part 2 introduces portfolio loans and for loss given default models. Part 3 details the calculation of the expected loss. Part 4 focusses on the conditional loss for a given systematic risk trajectory. Part 5 specifies the correlation structure between the systematic risk factors. Part 6 relates to the calibration of the risk factors using dynamic macro and micro correlations coming from climate models and climate analysts. Part 7 opens a perspective on reverse stress tests. The Annex gives some technical details on the recovery model.

This work was initiated by the association Green RWA (Risk Weighted Assets). It was written in collaboration with Jean-Baptiste Gaudemet, Anne Gruz and Olivier Vinciguerra, who contributed their financial and risk expertise, taking care of its application to a pilot-portfolio. It extends the model proposed in a first white paper published by Green RWA.

*CMAP, Ecole polytechnique, https://www.josselin-garnier.org; Lusenn, http://www.lusenn.com/; Green RWA, https://www.greenrwa.org/

1 Introduction

The goal of this paper is to assess the loss of a credit portfolio under stressed conditions. The portfolio comprises loans from a large number of borrowers, made up of different groups representing geographic regions, and/or economic sectors, and/or climate risk mitigation and adaptation strategies, and/or collateral types and which have different ratings at the initial time.

The expected loss can be expressed in terms of probability of default, exposure at default, and loss given default.

The unconditional migration matrices which express the probabilities for one borrower to move from one rating to another (until the ultimate rating which corresponds to default) during an arbitrary unit time interval, are supposed to be known. These migration matrices are unconditional in the sense that they are averaged over idiosyncratic and systematic risk factors.

The loss given default is based on a random recovery model that also depends on idiosyncratic and systematic risk factors. The correlation between default occurrence and recovery rate is only through the systematic risk factors. This model can be simplified in order to consider deterministic recovery rates imposed by the regulator.

The exposure at default (given default at a certain time) is explicit and independent from the migration and recovery processes. It can model various banking portfolio dynamics (flat, amortized...).

The systematic risk factors model economic, transition, and physical risks. They may be correlated and they influence all borrowers. The idiosyncratic risk factors are specific to each borrower.

The unexpected loss under stressed conditions is the $1-\alpha$-quantile of the conditional loss given the systematic risk factors (say, $1-\alpha = 99.9\%$). This approach follows the Basel IRB ASRF model (Internal Rating Based, Asymptotic Single Risk Factor [8, 9]) which calculates the loss conditional to a single systematic economic factor downturn. It is inspired by the Multi-Factor Merton model [6] and it extends the model proposed in the first white paper [3].

The paper is organized as follows. In Section 2 we present the model for the default occurrence and the recovery rate. In Section 3 we give the expected loss of a portfolio. In Section 4 we give the conditional loss of a portfolio under stressed conditions. We specify the models for the systematic risk factors in Section 5 and for the loading factors in Section 6.

2 The model for default and recovery

Each borrower belongs to a group and has a rating. We assume that:

- There are G groups. A group can represent a geographic region, and/or an economic sector, and/or a climate risk mitigation and adaptation strategy[1] and/or a collateral type.
- There are K rating levels $\{1, \ldots, K\}$. The rating K corresponds to default.

The time is discretized as integers $t = 0, \ldots, T$, where $t = 0$ is present and $t = T$ is the time horizon of the stress test analysis.

We consider a structural model such as the one proposed by Merton [5], and considerably extended in the literature [4, 9], where a borrower defaults, when the value of its asset falls below a threshold level. Default correlation is introduced by assuming that the assets of the borrowers are correlated stochastic processes. We adopt the Gaussian copula model. After normalization, we can write the log asset value at time t of the q-th borrower that belongs to the g-th group and has rating i at time $t-1$ in the form

$$X_t^{(q)} = \boldsymbol{a}_{g,i,t} \cdot \boldsymbol{Z}_t + \sqrt{1 - \boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{a}_{g,i,t}}\varepsilon_t^{(q)} \tag{1}$$

where

- The random vector $\boldsymbol{Z}_t$ contains the systematic (economic, physical, and transition) risk factors at time t. The vector $\boldsymbol{Z}_t$ is assumed to have multivariate normal distribution with mean $\mathbf{0}$ and correlation matrix $\mathbf{C}$. If the systematic risk factors are uncorrelated, then they are i.i.d. with standard normal distribution $\mathbf{C} = \mathbf{I}$. If the systematic risk factors are correlated, then this general model is necessary (see Section 5).

[1]Climate change mitigation consists of actions to lessen the magnitude or the rate of global warming and its related effects. This generally involves reductions in emissions of greenhouse gases (GHGs). Climate change adaptation consists of incremental adaptation actions where the central aim is to maintain the essence and integrity of a system or of transformational adaptation actions that change the fundamental attributes of a system in response to climate change and its impacts.

- The vectors $\boldsymbol{a}_{g,i,t}$ are the factor loadings (the correlations between the systematic risk factors and the assets) for the borrowers that belong to group g and have rating i at time $t-1$ (see Section 6).
- The idiosyncratic factors $\varepsilon_t^{(q)}$ are independent and identically distributed (i.i.d.) with standard normal distribution and independent from $\boldsymbol{Z}_t$; they model the risk specific to each borrower.

The borrower defaults when its normalized log asset value falls below an unconditional threshold value that corresponds to the unconditional probability of default of its group and rating. The loss is then denoted by $l_t^{(q)}$ and has the form

$$l_t^{(q)} = \mathrm{EAD}_t^{(q)} \big[1 - \mathrm{RR}_t^{(q)}\big], \tag{2}$$

where

- $\mathrm{EAD}_t^{(q)}$ is the exposure (the total balance owed by the borrower at time of default) of the q-th borrower given default at time t,
- $\mathrm{RR}_t^{(q)}$ is the recovery rate (the proportion of the exposure that is recovered by way of liquidation of collateral and other resolution or post-default collection actions) of the q-th borrower given default at time t. The loss given default (the proportion of the exposure that is lost if the borrower defaults) is $1 - \mathrm{RR}_t^{(q)}$.

The exposure at default is explicit and independent from the asset and recovery processes. It is determined by the principal and the amortization profile of the loan. For instance, for an amortizing loan with principal $K^{(q)}$, maturity $T^{(q)}$, interest rate $r^{(q)}$ and equal payments, we have

$$\mathrm{EAD}_t^{(q)} = K^{(q)} \frac{(1+r^{(q)})^{T^{(q)}} - (1+r^{(q)})^t}{(1+r^{(q)})^{T^{(q)}} - 1} \mathbf{1}_{t \leq T^{(q)}}. \tag{3}$$

Note that $T^{(q)}$ can be larger or smaller than T, which means there is no constraint on the distribution of the loan maturities.

The recovery rate of the q-th borrower that belongs to group g and has rating i at time $t-1$ has the general form inspired from [1]

$$\mathrm{RR}_t^{(q)} = \Phi\Big(\mu_{g,i,t} + \sigma_{g,i,t}\big(\boldsymbol{b}_{g,i,t} \cdot \boldsymbol{Z}_t + \sqrt{1 - \boldsymbol{b}_{g,i,t} \cdot \mathbf{C}\boldsymbol{b}_{g,i,t}}\,\tilde{\varepsilon}_t^{(q)}\big)\Big). \tag{4}$$

The recovery rate $\mathrm{RR}_t^{(q)}$ can be influenced by the same systematic risk factors $\boldsymbol{Z}_t$ as the assets:

- The vectors $\boldsymbol{b}_{g,i,t}$ are the factor loadings (the correlations between the systematic risk factors and the recovery rates). We may take $\boldsymbol{b}_{g,i,t} = \lambda_{g,i,t}\boldsymbol{a}_{g,i,t}$ in order to simplify the model, which means that the collateral is of the same type as the principal, and then $\lambda_{g,i,t}$ determines the dependence between the default occurrence and the recovery rate. The collateral, however, may be taken of a different type from the principal and then $\boldsymbol{b}_{g,i,t}$ is not collinear to $\boldsymbol{a}_{g,i,t}$.

- The idiosyncratic factors $\tilde{\varepsilon}_t^{(q)}$ are i.i.d. with standard normal distribution and independent from $\boldsymbol{Z}_t$ and $\varepsilon_t^{(q)}$; they model the risk affecting the recovery rate specific to each borrower.

- The parameters $\mu_{g,i,t}$ and $\sigma_{g,i,t}$ make it possible to fit observed distributions of recovery rates given default. Note that the distribution of $\mathrm{RR}_t^{(q)}$ given default is the distribution of $\mathrm{RR}_t^{(q)}$ given $X_t^{(q)}$ is below the threshold value corresponding to default, as explained in Appendix A.2.

- In the simple case when the recovery rates are deterministic and equal to $\mathrm{RR}_{g,i,t}$ that depend only on the group g, the rating i, and time t, we have
$$\mu_{g,i,t} = \Phi^{-1}(\mathrm{RR}_{g,i,t}) \text{ and } \sigma_{g,i,t} = 0, \tag{5}$$
and $\boldsymbol{b}_{g,i,t}$ plays no role (we may take $\boldsymbol{b}_{g,i,t} = \mathbf{0}$).

- If $\boldsymbol{b}_{g,i,t} = \mathbf{0}$, then the recovery rates are random but only through the idiosyncratic risk factor. The recovery rate $\mathrm{RR}_t^{(q)} = \Phi\big(\mu_{g,i,t}+\sigma_{g,i,t}\tilde{\varepsilon}_t^{(q)}\big)$ is independent from $X_t^{(q)}$ and the distribution of $\mathrm{RR}_t^{(q)}$ given default is of the form $\mathbb{P}(\mathrm{RR}_t^{(q)} \leq r|\text{default}) = \Phi[(\Phi^{-1}(r) - \mu_{g,i,t})/\sigma_{g,i,t}]$.

- If $\boldsymbol{b}_{g,i,t} \cdot \mathbf{C}\boldsymbol{b}_{g,i,t} = 1$, then the recovery rates are random but only through the systematic risk factors. We have $\mathrm{RR}_t^{(q)} = \Phi\big(\mu_{g,i,t} + \sigma_{g,i,t}\boldsymbol{b}_{g,i,t} \cdot \boldsymbol{Z}_t\big)$ is correlated to $X_t^{(q)}$ and the distribution of $\mathrm{RR}_t^{(q)}$ given default is complex (see Appendix A.2).

- The choice of the function Φ (the cdf of the standard normal distribution) is convenient to get closed form expressions and it allows (with the two parameters $\mu_{g,i,t}$ and $\sigma_{g,i,t}$) to match a large diversity of recovery rate distributions.

Note that, in this random recovery model, the loss given default and the default occurrence are correlated only through the systematic risk factors.

3 Expected loss

The expected loss is

$$L^{\mathrm{e}} = \sum_{t=1}^{T} L_t^{\mathrm{e}}, \tag{6}$$

$$L_1^{\mathrm{e}} = \sum_{g=1}^{G} \sum_{i=1}^{K-1} (\mathbf{M}_{g,1})_{iK} \mathrm{LGD}_{g,i,1} \mathrm{EAD}_{g,i,1}, \tag{7}$$

$$L_t^{\mathrm{e}} = \sum_{g=1}^{G} \sum_{i,j=1}^{K-1} \big(\mathbf{M}_{g,1} \cdots \mathbf{M}_{g,t-1}\big)_{ij} (\mathbf{M}_{g,t})_{jK} \mathrm{LGD}_{g,j,t} \mathrm{EAD}_{g,i,t}, \tag{8}$$

for $t \geq 2$, where:

- L_t^{e} is the expected loss due to the defaults that occur at time t. The first term L_1^{e} in (6) is the loss due to the borrowers in group g and with rating i at time 0 who default at time 1. The t-th term L_t^{e} is the loss due to the borrowers in group g and with rating i at time 0 who default at time t, it is decomposed on all possible rating $j < K$ that the borrowers may have at time $t-1$.

- $\mathrm{EAD}_{g,i,t}$ is the total exposure at default, given default at time t due to the borrowers in group g and with initial rating i:

$$\mathrm{EAD}_{g,i,t} = \sum_{q=1}^{N} \mathrm{EAD}_t^{(q)} \mathbf{1}_{q\text{-th borrower is in group } g \text{ and has initial rating } i}. \tag{9}$$

  $\mathrm{EAD}_{g,i,t}$ can be seen as the maximal loss at time t from the borrowers in group g and with initial rating i (in the worst case scenario when they all default at time t with zero recovery rate).

- $\mathbf{M}_{g,t}$ is the unconditional migration matrix (of size $K \times K$) at time t; $(\mathbf{M}_{g,t})_{ij}$ is the probability for a borrower in group g and with rating i at time $t-1$ to migrate to rating j at time t (see Subsection 3.1). In particular the i-th entry of the last column $(\mathbf{M}_{g,t})_{iK}$ gives the probability of default at time t for a borrower in group g and with rating $i \in \{1, \ldots, K-1\}$ at time $t-1$.
  $\big(\mathbf{M}_{g,1} \cdots \mathbf{M}_{g,t-1}\big)_{ij} (\mathbf{M}_{g,t})_{jK}$ can be interpreted as the probability that a borrower in group g and with rating i at time zero has rating j at time $t-1$ and defaults at time t.

- $\mathrm{LGD}_{g,j,t}$ is the average Loss Given Default for a borrower in group g and with rating j at time $t-1$ who defaults at time t (its rating jumps from j to K) (see Subsection 3.2).

Remark. The framework proposed in this paper can be used when new loans are added to the portfolio at different times. Indeed, if new loans are added at time $t_0 > 0$, then they can be incorporated into the model by creating new groups g' which are such that the migration matrices $\mathbf{M}_{g',t}$ are equal to the identity matrix $\mathbf{I}$ for $t \le t_0$. This would make it possible to address various dynamic balance sheet strategies (where the composition or risk profile of the portfolio is allowed to vary over the stress test horizon), as long as these strategies depend only on the expected losses of the different groups.

Remark. The framework proposed in this paper can be used when the portfolio amortizes and adds new loans in a balanced way: at any time, the amortization of the previous loans is compensated for by the addition of fresh loans. More precisely, let us address the case where, for each group g:
- the initial rating profile and the rating profile of the new loans is described by the vector $\boldsymbol{w}_g$: $w_{g,i}$ is the proportion of loans with rating i in the group g at time 0, with $w_{g,K} = 0$ and $\sum_{i=1}^{K-1} w_{g,i} = 1$.
- a fraction $1 - \kappa_g$ of loans is amortized every unit time and a fraction κ_g of new loans with the rating profile $\boldsymbol{w}_g$ is added every unit time.
- the exposure at default for the group is kept constant at EAD_g (it is of course possible to consider a time-dependent evolution).
This situation can be modelled in the proposed framework, provided we use the updated migration matrices

$$\mathbf{M}^{\boldsymbol{w}}_{g,t} = (1-\kappa_g)\mathbf{M}_{g,t} + \kappa_g \mathbf{M}^{\boldsymbol{w}}_g, \tag{10}$$

$$\mathbf{M}^{\boldsymbol{w}}_g = \mathbf{1}\boldsymbol{w}_g^T = \begin{pmatrix} w_{g,1} & \dots & w_{g,K-1} & 0 \\ \vdots & & \vdots & \vdots \\ w_{g,1} & \dots & w_{g,K-1} & 0 \end{pmatrix}, \tag{11}$$

where $\mathbf{1}$ is the K-dimensional vector full of ones. The expected loss is then given by (6-8) with the matrices $\mathbf{M}^{\boldsymbol{w}}_{g,t}$ instead of $\mathbf{M}_{g,t}$ and $\mathrm{EAD}^{\boldsymbol{w}}_{g,i,t} = \mathrm{EAD}_g w_{g,i}$. The unexpected loss that we address in the next section is then given by (20-22) with the conditional matrices $\mathbf{M}^{\boldsymbol{w}}_{g,t}(\boldsymbol{Z}_t) = (1-\kappa_g)\mathbf{M}_{g,t}(\boldsymbol{Z}_t) + \kappa_g \mathbf{M}^{\boldsymbol{w}}_g$ where $\mathbf{M}_{g,t}(\boldsymbol{Z}_t)$ is given by (28) and $\mathbf{M}^{\boldsymbol{w}}_g$ is given by (11).

3.1 Unconditional migration matrices

The rating K corresponds to default, it is an absorbing state. The $K \times K$ matrix $\mathbf{M}_{g,t}$ has non negative entries, it satisfies $\sum_{j=1}^{K} (\mathbf{M}_{g,t})_{ij} = 1$ and $(\mathbf{M}_{g,t})_{KK} = 1$.

$(\mathbf{M}_{g,t})_{ij}$ is the probability for a borrower in group g and with rating i at time $t-1$ to migrate to rating j at time t. A borrower in group g with rating i at time $t-1$ will migrate to a rating in the interval $[j, K]$ if its normalized log asset value falls below the unconditional threshold value $z_{g,t,ij}$. The unconditional distribution of the normalized log asset value (1) of a borrower is standard normal,

$$\mathbb{P}\big(X_t^{(q)} \le z_{g,t,ij}\big) = \Phi(z_{g,t,ij}), \tag{12}$$

so the unconditional threshold values are given in terms of quantiles of the standard normal distribution:

$$z_{g,t,ij} = \Phi^{-1}\Big(\sum_{j'=j}^{K} (\mathbf{M}_{g,t})_{ij'}\Big). \tag{13}$$

Note that:
- $z_{g,t,i1} = +\infty$ for all $i \le K$ because $\sum_{j'=1}^{K} (\mathbf{M}_{g,t})_{ij'} = 1$.
- The term $z_{g,t,iK}$ is the unconditional threshold value that corresponds to the unconditional probability of default at time t for a borrower in group g with rating i at time $t-1$.
- $z_{g,t,Kj} = +\infty$ for $j \le K$ because $(\mathbf{M}_{g,t})_{KK} = 1$.

3.2 Average loss given default

By (2) the average Loss Given Default for the borrowers in group g with rating i at time $t-1$ that default at time t is

$$\mathrm{LGD}_{g,i,t} = \mathbb{E}\big[1 - \mathrm{RR}_t^{(q)} | X_t^{(q)} \le z_{g,t,iK}\big], \tag{14}$$

because the event "$X_t^{(q)} \le z_{g,t,iK}$" corresponds to default for the q-th borrower (which belongs to group g and has rating i at time $t-1$). As shown in Appendix A.2, the average Loss Given Default for the borrowers from group g and with rating i at time $t-1$ who default at time t depends on the rating i:

$$\mathrm{LGD}_{g,i,t} = 1 - \frac{1}{(\mathbf{M}_{g,t})_{iK}} \Phi_2\Big(\frac{\mu_{g,i,t}}{\sqrt{1+\sigma_{g,i,t}^2}}, z_{g,t,iK}; \frac{-\rho_{g,i,t}\sigma_{g,i,t}}{\sqrt{1+\sigma_{g,i,t}^2}}\Big), \tag{15}$$

$$\rho_{g,i,t} = \boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{b}_{g,i,t}, \tag{16}$$

where $\Phi_2(\cdot,\cdot;\rho)$ is the bivariate cumulative Gaussian distribution with correlation ρ. $\rho_{g,i,t}$ is the linear correlation coefficient between the normalized log asset value $X_t^{(q)}$ and $\Phi^{-1}(\mathrm{RR}_t^{(q)})$. The Kendall's Tau between the normalized log asset value $X_t^{(q)}$ and the recovery rate $\mathrm{RR}_t^{(q)}$ is (see Appendix A.1):

$$\tau(X_t^{(q)},\mathrm{RR}_t^{(q)}) = \frac{2}{\pi}\arcsin\bigl(\rho_{g,i,t}\bigr). \tag{17}$$

Eq. (15) (and also (17)) can be used to calibrate the parameters of the recovery model from default swap market data. More elaborate moment matching can be used because it is also possible to express, in simple closed forms, the moments $\mathbb{E}\bigl[\Phi^{-1}(\mathrm{RR}_t^{(q)})^n|X_t^{(q)} \le z_{g,t,iK}\bigr]$ of the recovery rate for a borrower with rating i at time $t-1$ who defaults at time t (see Appendix A.2).

Of course:

- If the recovery rate and the default occurrence are independent (i.e. if $\rho_{g,i,t}=0$), then

$$\mathrm{LGD}_{g,i,t} = 1-\Phi\Bigl(\frac{\mu_{g,i,t}}{\sqrt{1+\sigma_{g,i,t}^2}}\Bigr)$$

  is equal to one minus the expected recovery rate for a borrower that belongs to group g and has rating i at time $t-1$:

$$\mathrm{LGD}^{\mathrm{e}}_{g,i,t} = \mathbb{E}[1-\mathrm{RR}_t^{(q)}].$$

- If the recovery rate is deterministic and equal to $\mathrm{RR}_{g,i,t}$ (i.e. if $\sigma_{g,i,t}=0$) then

$$\mathrm{LGD}^{\mathrm{e}}_{g,i,t} = 1-\mathrm{RR}_{g,i,t}. \tag{18}$$

4 Conditional loss

We assume that the portfolio is large. More exactly, we assume that the portfolio contains a large number N of loans without it being dominated by a few loans much larger than the rest. This hypothesis can be formulated as the non-concentration condition

$$\frac{\sum_{q=1}^N(\mathrm{EAD}_t^{(q)})^2}{\bigl[\sum_{q=1}^N(\mathrm{EAD}_t^{(q)})\bigr]^2} \overset{N\to\infty}{\longrightarrow} 0. \tag{19}$$

This hypothesis implies that the idiosyncratic risks are diversified, but not the systematic risks. Then, by the law of large numbers, the conditional loss, given a trajectory $\mathbf{Z}$ of the systematic (economic, physical and transition) risk factors, is

$$L(\mathbf{Z}) = \sum_{t=1}^{T} L_t(\mathbf{Z}), \tag{20}$$

$$L_1(\mathbf{Z}) = \sum_{g=1}^{G} \sum_{i=1}^{K-1} (\mathbf{M}_{g,1}(\boldsymbol{Z}_1))_{iK} \mathrm{LGD}_{g,i,1}(\boldsymbol{Z}_1) \mathrm{EAD}_{g,i,1}, \tag{21}$$

$$L_t(\mathbf{Z}) = \sum_{g=1}^{G} \sum_{i,j=1}^{K-1} \left(\mathbf{M}_{g,1}(\boldsymbol{Z}_1) \cdots \mathbf{M}_{g,t-1}(\boldsymbol{Z}_{t-1})\right)_{ij} (\mathbf{M}_{g,t}(\boldsymbol{Z}_t))_{jK} \mathrm{LGD}_{g,j,t}(\boldsymbol{Z}_t) \mathrm{EAD}_{g,i,t}, \tag{22}$$

for $t \geq 2$, where:

- $L_t(\mathbf{Z})$ is the conditional loss due to the defaults that occur at time t.
- $\mathbf{Z} = (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_T)$ is the trajectory of the systematic risk factors.
- $\mathrm{EAD}_{g,i,t}$ is the total exposure at default (9) given default at time t due to the borrowers in group g and with initial rating i.
- $\mathbf{M}_{g,t}(\boldsymbol{Z}_t)$ is the conditional migration matrix (of size $K \times K$); $(\mathbf{M}_{g,t}(\boldsymbol{Z}_t))_{ij}$ is the probability for a borrower in group g and with rating i at time $t-1$ to migrate to rating j at time t, given the systematic risk factors $\boldsymbol{Z}_t$ during this period (see Eq. (28)).
- $\mathrm{LGD}_{g,i,t}(\boldsymbol{Z}_t)$ is the conditional Loss Given Default for a borrower in group g and with rating i at time $t-1$ who defaults at time t (its rating jumps from i to K), given the systematic risk factors $\boldsymbol{Z}_t$ during this period (see Eq. (29)).

If the non-concentration condition (19) is not fulfilled, then granularity adjustment is necessary to take into account that the portfolio may carry some undiversified idiosyncratic risk [9, 7, 2].

The conditional loss $L(\mathbf{Z})$ is a deterministic function of the trajectory $\mathbf{Z}$. In the next subsections we present closed form expressions for the conditional migration matrix $\mathbf{M}_{g,t}(\boldsymbol{Z}_t)$ and the conditional Loss Given Default $\mathrm{LGD}_{g,i,t}(\boldsymbol{Z}_t)$. As a result we have closed form expressions for the conditional loss $L(\mathbf{Z})$ and the conditional partial losses $L_t(\mathbf{Z})$.

Given a distribution for the process $\mathbf{Z}$, the conditional loss in stressed conditions $L_{\text{stress}}^{1-\alpha}$ is the $1-\alpha$-quantile of $L(\mathbf{Z})$:

$$\mathbb{P}(L(\mathbf{Z}) \leq L_{\text{stress}}^{1-\alpha}) = 1-\alpha. \tag{23}$$

with typically $\alpha = 10^{-3}$ ($1-\alpha = 99.9\%$) or $\alpha = 10^{-3}T$. A straightforward method to estimate this quantile is a Monte Carlo method with a sample size $(\mathbf{Z}^{(k)})_{k=1}^{N_{\text{MC}}}$ of the order of $N_{\text{MC}} = 100/\alpha$. The estimator is the empirical $1-\alpha$-quantile of the sample $(L(\mathbf{Z}^{(k)}))_{k=1}^{N_{\text{MC}}}$. Variance reduction techniques, (such as importance sampling), can be implemented to reduce the required sample size.

Given a distribution for the process $\mathbf{Z}$, the conditional loss in stressed conditions $L_{t,\text{stress}}^{1-\alpha}$ during the t-th period, (the t-th year when the time unit is one year), is the $1-\alpha$-quantile of $L_t(\mathbf{Z})$:

$$\mathbb{P}(L_t(\mathbf{Z}) \leq L_{t,\text{stress}}^{1-\alpha}) = 1-\alpha. \tag{24}$$

Note that $L_t(\mathbf{Z})$, $t = 1, \ldots, T$ are correlated and are, of course, correlated with $L(\mathbf{Z})$ since $L(\mathbf{Z}) = \sum_{t=1}^{T} L_t(\mathbf{Z})$. We have:

$$\sum_{t=1}^{T} L_{t,\text{stress}}^{1-\alpha/T} \geq L_{\text{stress}}^{1-\alpha}. \tag{25}$$

Proof. We have for any ℓ_t

$$\begin{aligned}
\mathbb{P}\Big(L(\mathbf{Z}) \leq \sum_{t=1}^{T} \ell_t\Big) &= \mathbb{P}\Big(\sum_{t=1}^{T} L_t(\mathbf{Z}) \leq \sum_{t=1}^{T} \ell_t\Big) \\
&\geq \mathbb{P}\Big(\bigcap_{t=1}^{T} \big\{L_t(\mathbf{Z}) \leq \ell_t\big\}\Big) \\
&= 1 - \mathbb{P}\Big(\bigcup_{t=1}^{T} \big\{L_t(\mathbf{Z}) > \ell_t\big\}\Big) \\
&\geq 1 - \sum_{t=1}^{T} \mathbb{P}\big(L_t(\mathbf{Z}) > \ell_t\big).
\end{aligned}$$

By taking $\ell_t = L_{t,\text{stress}}^{1-\alpha/T}$ we get

$$\mathbb{P}\Big(L(\mathbf{Z}) \leq \sum_{t=1}^{T} L_{t,\text{stress}}^{1-\alpha/T}\Big) \geq 1 - T\frac{\alpha}{T} = 1-\alpha,$$

which gives the desired result. □

The regulatory capital charge K_t at time t for the portfolio is

$$K_t = L^{1-\alpha}_{\mathrm{stress},t} - L^{\mathrm{e}}_t, \tag{26}$$

with the expected loss L^{e}_t given by (8) and the unexpected loss $L^{1-\alpha}_{\mathrm{stress},t}$ given by (24). It can be multiplied by a maturity adjustment factor, given by the foundation IRB model when the unit time is one year for instance. It could also be possible to compute an average capital charge, that would be K/T where $K = L^{1-\alpha T}_{\mathrm{stress}} - L^{\mathrm{e}}$, with the expected loss L^{e} given by (6) and the unexpected loss $L^{1-\alpha T}_{\mathrm{stress}}$ given by (23).

4.1 Conditional migration matrices

We assume that we know the unconditional migration matrices $\mathbf{M}_{g,t}$ for each group g.

Given $\boldsymbol{Z}_t$, a borrower in group g with rating i at time $t-1$ will migrate to a rating in the interval $[j, K]$ at time t if its normalized log asset value (given $\boldsymbol{Z}_t$) falls below the threshold $z_{g,t,ij}$. This event has probability

$$\mathbb{P}\left(X^{(q)}_t \le z_{g,t,ij} | \boldsymbol{Z}_t\right) = \Phi\left(\frac{z_{g,t,ij} - \boldsymbol{a}_{g,i,t} \cdot \boldsymbol{Z}_t}{\sqrt{1 - \boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{a}_{g,i,t}}}\right). \tag{27}$$

As a consequence, the conditional migration matrix $\mathbf{M}_{g,t}(\boldsymbol{Z}_t)$ is given by

$$(\mathbf{M}_{g,t}(\boldsymbol{Z}_t))_{ij} = \begin{cases} 1 - \Phi\left(\frac{z_{g,t,i2} - \boldsymbol{a}_{g,i,t} \cdot \boldsymbol{Z}_t}{\sqrt{1 - \boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{a}_{g,i,t}}}\right), & \text{if } j = 1, \\ \Phi\left(\frac{z_{g,t,ij} - \boldsymbol{a}_{g,i,t} \cdot \boldsymbol{Z}_t}{\sqrt{1 - \boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{a}_{g,i,t}}}\right) \\ \qquad -\Phi\left(\frac{z_{g,t,ij+1} - \boldsymbol{a}_{g,i,t} \cdot \boldsymbol{Z}_t}{\sqrt{1 - \boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{a}_{g,i,t}}}\right), & \text{if } 2 \le j \le K-1, \\ \Phi\left(\frac{z_{g,t,iK} - \boldsymbol{a}_{g,i,t} \cdot \boldsymbol{Z}_t}{\sqrt{1 - \boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{a}_{g,i,t}}}\right), & \text{if } j = K. \end{cases} \tag{28}$$

The entries of the last column $(\mathbf{M}_{g,t}(\boldsymbol{Z}_t))_{iK}$ gives the conditional probability of default for a borrower in group g with rating $i = 1, \ldots, K-1$ at time $t-1$.

4.2 Conditional loss given default

The particular form of the loss (2) makes it possible to give a simple closed form formula for the conditional Loss Given Default $\mathrm{LGD}_{g,i,t}(\boldsymbol{Z}_t)$:

$$\mathrm{LGD}_{g,i,t}(\boldsymbol{Z}_t) = 1 - \Phi\Big(\frac{\mu_{g,i,t} + \sigma_{g,i,t}\boldsymbol{b}_{g,i,t} \cdot \boldsymbol{Z}_t}{\sqrt{1 + \sigma^2_{g,i,t}(1 - \boldsymbol{b}_{g,i,t} \cdot \mathbf{C}\boldsymbol{b}_{g,i,t})}}\Big), \tag{29}$$

Of course, if the recovery rate is deterministic and equal to $\mathrm{RR}_{g,i,t}$ (i.e. if $\sigma_{g,i,t} = 0$) then

$$\mathrm{LGD}_{g,i,t}(\boldsymbol{Z}_t) = 1 - \mathrm{RR}_{g,i,t}. \tag{30}$$

4.3 An explicit and simple case

If $T = 1$, $G = 1$, $\mathbf{C} = \mathbf{I}$, $\sigma_{1,i,1} = 0$ (the LGD is deterministic), and $\boldsymbol{a}_{1,i,1} = \boldsymbol{a}_1$ (all borrowers have the same exposition with respect to the systematic risks whatever their inital rating), then we can get a closed form expression for the conditional loss in stressed conditions $L^{1-\alpha}_{\text{stress}}$. Indeed we have

$$\begin{aligned} L(\mathbf{Z}) &= \sum_{i=1}^{K-1} (\mathbf{M}_{1,1}(\boldsymbol{Z}_1))_{iK}(1 - \mathrm{RR}_{1,i,1})\mathrm{EAD}_{1,i,1} \\ &= \mathcal{L}(\boldsymbol{a}_1 \cdot \boldsymbol{Z}_1), \end{aligned} \tag{31}$$

where

$$\mathcal{L}(z) = \sum_{i=1}^{K-1} \Phi\Big(\frac{z_{1,1,iK} - z}{\sqrt{1 - \|\boldsymbol{a}_1\|^2}}\Big)(1 - \mathrm{RR}_{1,i,1})\mathrm{EAD}_{1,i,1}. \tag{32}$$

The function $z \mapsto \mathcal{L}(z)$ is decreasing, so we have for any ℓ:

$$\begin{aligned} \mathbb{P}\big(L(\mathbf{Z}) \le \ell\big) &= \mathbb{P}\big(\boldsymbol{a}_1 \cdot \boldsymbol{Z}_1 \ge \mathcal{L}^{-1}(\ell)\big) \\ &= 1 - \Phi\Big(\frac{\mathcal{L}^{-1}(\ell)}{\|\boldsymbol{a}_1\|}\Big) = \Phi\Big(-\frac{\mathcal{L}^{-1}(\ell)}{\|\boldsymbol{a}_1\|}\Big), \end{aligned} \tag{33}$$

because the random variable $\boldsymbol{a}_1 \cdot \boldsymbol{Z}_1$, has distribution $\mathcal{N}(0, \|\boldsymbol{a}_1\|^2)$. The conditional loss in stressed conditions (23) is therefore (with $\alpha = 10^{-3}$)

$$\begin{aligned} L^{0.999}_{\text{stress}} &= \mathcal{L}\big(-\Phi^{-1}(0.999)\|\boldsymbol{a}_1\|\big) \\ &= \sum_{i=1}^{K-1} \Phi\Big(\frac{z_{1,1,iK} + \Phi^{-1}(0.999)\|\boldsymbol{a}_1\|}{\sqrt{1 - \|\boldsymbol{a}_1\|^2}}\Big)(1 - \mathrm{RR}_{1,i,1})\mathrm{EAD}_{1,i,1}. \end{aligned}$$

If $K = 2$ (which means that there are only two ratings: "healthy" and "default"), then $z_{1,1,12} = \Phi^{-1}(\mathrm{PD})$ where PD is the unconditional probability of default and we recover the formula of the first version of the white paper

$$L^{0.999}_{\text{stress}} = \Phi\Big(\frac{\Phi^{-1}(\mathrm{PD}) + \Phi^{-1}(0.999)\|\boldsymbol{a}_1\|}{\sqrt{1-\|\boldsymbol{a}_1\|^2}}\Big)(1 - \mathrm{RR}_{1,1,1})\mathrm{EAD}_{1,1,1}. \tag{34}$$

5 Model for the systematic risk factors

The vector $\boldsymbol{Z}_t$ contains d systematic risk factors.

5.1 Independent risk factors

Here we consider models in which $\boldsymbol{Z}_t$ has i.i.d. entries with standard normal distribution, i.e. $\mathbf{C} = \mathbf{I}$.

When one wishes to study economic systematic risk, one usually uses a one-factor model. In this model the Z_t are i.i.d. with standard normal distribution.

Here we want to study economic, physical, and transition systematic risks. One can therefore consider a three-factor model $\boldsymbol{Z}_t = (Z_{t,1}, Z_{t,2}, Z_{t,3})$ where $Z_{t,1}$ is the economic risk factor, $Z_{t,2}$ is the transition risk factor, and $Z_{t,3}$ is the physical risk factor. We can take them independently with standard normal distribution.

We can also make the model more complex by considering several independent physical risk factors, one per region. If we want to model groups that are exposed to only one regional physical risk, then we would need to index the group as follows: $g = (e, r)$, where $e = 1, \dots, E$ is the index of the non-geographical sector (economic sector and/or climate risk mitigation and adaptation strategy and/or collateral type), and $r = 1, \dots, R$ the index of the geographical region, so that there are $G = ER$ groups in total. The vector $\boldsymbol{Z}_t$ would then be of the form $\boldsymbol{Z}_t = (Z_{t,j})_{j=1}^{2+R}$ where $Z_{t,1}$ is the economic risk factor, $Z_{t,2}$ is the transition risk factor, and $Z_{t,(2+r)}$ is the physical risk factor of the r-th region, $r = 1, \dots, R$. The factor loadings would then be of the form $\boldsymbol{a}_{e,r,i,t} = (a_{e,r,i,t,j})_{j=1}^{2+R}$, where $a_{e,r,i,t,1}$ is the factor loading associated to the economic risk at time t of a borrower with rating i in non-geographical sector e and region r, $a_{e,r,i,t,2}$ is the factor loading associated to the transition risk for such a borrower, $a_{e,r,i,t,2+r}$ is the factor loading associated to the physical risk of the r-th region for such a borrower, the factor loadings associated to the physical risks of the other regions are zero: $a_{e,r,i,t,2+r'} = 0$ for $r' \neq r$. We can,for instance, also introduce other groups that are exposed to several regional physical risks simultaneously.

5.2 Correlated risk factors

Here we consider models in which $\boldsymbol{Z}_t$ has multivariate normal distribution with mean $\mathbf{0}$ and correlation matrix $\mathbf{C}$. These models are necessary if we want to model correlations between some systematic risk factors.

We may think at an example where $\boldsymbol{Z}_t = (Z_{t,j})_{j=1}^{2+R}$, $Z_{t,1}$ is the economic risk factor, $Z_{t,2}$ is the transition risk factor, and $Z_{t,2+r}$ is the physical risk factor of the r-th region, $r = 1, \ldots, R$:

$$\boldsymbol{Z}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{C}) \text{ i.i.d.}, \tag{35}$$

with

$$\mathbf{C} = \begin{pmatrix} 1 & -\rho & 0 & 0 & \cdots & 0 \\ -\rho & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \rho_o & \cdots & \rho_o \\ 0 & 0 & \rho_o & 1 & \cdots & \rho_o \\ 0 & 0 & \rho_o & \rho_o & \ddots & \rho_o \\ 0 & 0 & \rho_o & \rho_o & \rho_o & 1 \end{pmatrix}, \tag{36}$$

which means that:
1) the physical risks of different geographical regions are positively correlated ($\rho_o \in (0,1)$) and independent from the economic and transition risks,
2) the transition risk is negatively correlated with the economic risk ($\rho \in (0,1)$). This comes from the observation that an economic downturn may involve a reduction in emissions of greenhouse gases.
The covariance matrix (36) can be made more complex, for instance, if correlations between physical risks in different regions are known.

6 Model for the loading factors

6.1 The model with a unique systematic risk factor

Under the foundation IRB (Internal Rating Based) approach:
- The time unit is one year.
- The LGD model is deterministic and imposed by the regulator. That is to say, $\text{LGD}_{g,i,t}$ are given by $\text{LGD}^{\text{reg}}_{g,i}$ that do not depend on t, but that depend on the group g and the rating i before default.
- The EAD model is deterministic and determined by the loan composition

of the portfolio.
- The unconditional migration matrices $\mathbf{M}_{g,t}$ are given by $\mathbf{M}_g^{\rm reg}$ that do not depend on t, but depend on the group g. The matrices $\mathbf{M}_g^{\rm reg}$ are typically estimated from historical data.
- The correlation model to economic risk (assumed to be the unique systematic risk factor) is determined by a formula that is imposed by the regulator and that is a function of the probability of default:

$$R_{g,i}^{\rm reg} = \mathcal{R}({\rm PD}_{g,i}^{\rm reg}), \tag{37}$$

$$\mathcal{R}({\rm PD}) = 0.12\frac{1-e^{-50{\rm PD}}}{1-e^{-50}} + 0.24\Big(1-\frac{1-e^{-50{\rm PD}}}{1-e^{-50}}\Big), \tag{38}$$

where ${\rm PD}_{g,i}^{\rm reg} = (\mathbf{M}_g^{\rm reg})_{iK}$ is the probability of default at time t of a borrower in group g and with rating i at time $t-1$. There is a unique systematic risk factor $Z_{t,1}$ and the loading factor $a_{g,i,t,1} = a_{g,i}^{\rm reg}$, which does not depend on t and is equal to

$$a_{g,i}^{\rm reg} = \sqrt{R_{g,i}^{\rm reg}}. \tag{39}$$

Under these hypotheses, the expected loss is given by (6):

$$\begin{aligned} L^{\rm e} &= \sum_{t=1}^T L_t^{\rm e}, \\ L_1^{\rm e} &= \sum_{g=1}^G \sum_{i=1}^{K-1} (\mathbf{M}_g^{\rm reg})_{iK} {\rm LGD}_{g,i}^{\rm reg} {\rm EAD}_{g,i,1}, \\ L_t^{\rm e} &= \sum_{g=1}^G \sum_{i,j=1}^{K-1} \big((\mathbf{M}_g^{\rm reg})^{t-1}\big)_{ij} \big(\mathbf{M}_g^{\rm reg}\big)_{jK} {\rm LGD}_{g,j}^{\rm reg} {\rm EAD}_{g,i,t}, \end{aligned}$$

for $t \geq 2$, and the conditional loss given a trajectory $\mathbf{Z} = (Z_{1,1}, \dots, Z_{T,1})$ of the economic risk factor is given by (20):

$$\begin{aligned} L(\mathbf{Z}) &= \sum_{t=1}^T L_t(\mathbf{Z}), \\ L_1(\mathbf{Z}) &= \sum_{g=1}^G \sum_{i=1}^{K-1} (\mathbf{M}_g^{\rm reg}(Z_{1,1}))_{iK} {\rm LGD}_{g,i}^{\rm reg} {\rm EAD}_{g,i,1}, \\ L_t(\mathbf{Z}) &= \sum_{g=1}^G \sum_{i,j=1}^{K-1} \big(\mathbf{M}_g^{\rm reg}(Z_{1,1}) \cdots \mathbf{M}_g^{\rm reg}(Z_{t-1,1})\big)_{ij} \big(\mathbf{M}_g^{\rm reg}(Z_{t,1})\big)_{jK} {\rm LGD}_{g,j}^{\rm reg} {\rm EAD}_{g,i,t}, \end{aligned}$$

for $t \geq 2$, where the conditional migration matrices are given by (28). The conditional loss in stressed conditions $L^{1-\alpha}_{\text{stress}}$ is the $1-\alpha$-quantile of $L(\mathbf{Z})$ when the $Z_{t,1}$ are independent and identically distributed with the standard normal distribution.

6.2 The model with multiple systematic risk factors

We need to extend the previous model to take into account transition and physical risks. We still assume that $\text{LGD}^{\text{reg}}_{g,i}$ and $\mathbf{M}^{\text{reg}}_g$ are given. We need to extend the correlation model and its relation to the loading factors $\boldsymbol{a}_{g,i,t}$ for the systematic risk factors $\boldsymbol{Z}_t = (Z_{t,j})_{j=1}^{R+2}$ described in Section 5.2.

We introduce the macro-correlation parameters $\boldsymbol{\zeta}_t = (\zeta_{t,j})_{j=1}^{R+2}$. They give the evolutions of the intensities of the $R+2$ systematic risk factors (economic, transition, physical divided into R regions). $\zeta_{t,1}$ is associated to the economic risk and assumed to be constant and equal to ζ_1. $\zeta_{t,2}$ and $\zeta_{t,2+r}$ are associated to the transition and physical risks and evolve in time. These parameters are relative to each other and should be expressed in the same "units". For instance, we may express all macro-correlation parameters in terms of GDP growth rates. ζ_1 can be the GDP growth rate involved by an economic downturn. $\zeta_{t,2}$ can be calibrated from the IPCC (Intergovernmental Panel on Climate Change) carbon emission pathway expressed in impact to GDP growth rate. $\zeta_{t,2+r}$ can be calibrated from the IPCC GDP growth rate assessment for the region r. Macro-economic and macro-climatic data can also be obtained from the Network for Greening the Financial System (NGFS) or the International Energy Agency (IEA). From now on we assume that $\boldsymbol{\zeta}_t$ is given.

We introduce the micro-correlation adjustment parameters $\alpha_{g,i,t,j}$. Each borrower in group g and with rating i at time $t-1$ has a micro-correlation adjustment parameter $\alpha_{g,i,t,j}$ to the j-th systematic risk factor. This micro-correlation parameter depends on the group. It may depend on the rating. It may be time-dependent in order to take into account mitigation and adaptation efforts by the borrowers. Note that a micro-correlation adjustment parameter can be negative (for instance, transition risk may favour an economic sector). From now on we assume that $\boldsymbol{\alpha}_{g,i,t} = (\alpha_{g,i,t,j})_{j=1}^{R+2}$ are given.

We introduce the correlation $R_{g,i,t}$. The correlation is the proportion of the variance of the normalized log asset value that is due to the systematic risks. Equivalently, $1 - R_{g,i,t}$ is the proportion of the variance of the normalized log asset value that is due to the idiosyncratic risk of a borrower.

From (1) it is related to the factor loadings through the relation:

$$R_{g,i,t} = \boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{a}_{g,i,t}. \tag{40}$$

6.2.1 *Approach 1 for the correlation model and factor loadings.*

We consider here that:
- the time unit is one year,
- the migration matrices $\mathbf{M}_{g,t}$ do not depend on t and are equal to $\mathbf{M}_g^{\rm reg}$,
- the correlation $R_{g,i,t}$ at any time t is determined by the regulator's formula which does not depend on t,
- the factor loadings $a_{g,i,t,j}$ are proportional to the product of the macro-correlation and micro-correlation adjustment parameters.

As a result, we have

$$R_{g,i,t} = R_{g,i}^{\rm reg}, \qquad R_{g,i}^{\rm reg} = \mathcal{R}({\rm PD}_{g,i}^{\rm reg}), \tag{41}$$

with ${\rm PD}_{g,i}^{\rm reg} = (\mathbf{M}_g^{\rm reg})_{iK}$, $\mathcal{R}$ defined by (38), and

$$a_{g,i,t,j} = \sqrt{R_{g,i}^{\rm reg}} \frac{\tilde{a}_{g,i,t,j}}{\sqrt{\tilde{\boldsymbol{a}}_{g,i,t} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,t}}}, \tag{42}$$

with

$$\tilde{a}_{g,i,t,j} = \alpha_{g,i,t,j}\zeta_{t,j}. \tag{43}$$

Proof. The factor loadings $a_{g,i,t,j}$ are proportional to the product $\tilde{a}_{g,i,t,j}$ of the macro-correlation and micro-correlation adjustment parameters. From (1) the factor loadings also satisfy $\boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{a}_{g,i,t} = R_{g,i}^{\rm reg}$. This imposes the form (42) of the factor loadings. □

Remark. Note that, in this approach, when the intensities $\zeta_{t,2}$ and/or $\zeta_{t,2+r}$ increase (compared to ζ_1 that is constant), then the correlation $R_{g,i,t}$ is not affected because it is determined by the regulator's formula, which depends only on the given unconditional migration matrices. The only effect of the increase of the intensities $\zeta_{t,2}$ and/or $\zeta_{t,2+r}$ is to modify the proportions of the economic and climate contributions to the constrained value of the correlation $R_{g,i}^{\rm reg}$. To sum-up, if the climatic risk intensities increase, then the economic risk intensity decays in order to maintain the same correlation. This may be problematic.

6.2.2 *Approach 2 for the correlation model and loading factors.*

We consider here that:
- the time unit is one year,

- the migration matrices $\mathbf{M}_{g,t}$ do not depend on t and are equal to $\mathbf{M}_g^{\text{reg}}$,
- the correlation $R_{g,i,1}$ at time 1 is determined by the regulator's formula, but this formula is updated at time $t \geq 2$ because, contrary to the economic risk, which is stationary, the physical and transition risks evolve in time.
- the factor loadings $a_{g,i,t,j}$ are proportional to the product of the macro-correlation and micro-correlation adjustment parameters.

As a result we have

$$R_{g,i,t} = \frac{\tilde{\boldsymbol{a}}_{g,i,t} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,t} R_{g,i}^{\text{reg}}}{\tilde{\boldsymbol{a}}_{g,i,t} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,t} R_{g,i}^{\text{reg}} + \tilde{\boldsymbol{a}}_{g,i,1} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,1}(1 - R_{g,i}^{\text{reg}})}, \tag{44}$$

and

$$a_{g,i,t,j} = \sqrt{R_{g,i}^{\text{reg}}} \frac{\tilde{a}_{g,i,t,j}}{\sqrt{\tilde{\boldsymbol{a}}_{g,i,t} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,t} R_{g,i}^{\text{reg}} + \tilde{\boldsymbol{a}}_{g,i,1} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,1}(1 - R_{g,i}^{\text{reg}})}}, \tag{45}$$

with

$$\tilde{a}_{g,i,t,j} = \alpha_{g,i,t,j}\zeta_{t,j}. \tag{46}$$

Proof. At time 1 (see the first approach) the normalized log asset value is given by

$$X_1^{(q)} = \boldsymbol{a}_{g,i}^{\text{reg}} \cdot \boldsymbol{Z}_1 + \sqrt{1 - \boldsymbol{a}_{g,i}^{\text{reg}} \cdot \mathbf{C}\boldsymbol{a}_{g,i}^{\text{reg}}}\,\varepsilon_1^{(q)}, \tag{47}$$

with

$$a_{g,i,j}^{\text{reg}} = \frac{\sqrt{R_{g,i}^{\text{reg}}}\,\tilde{a}_{g,i,1,j}}{\sqrt{\tilde{\boldsymbol{a}}_{g,i,1} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,1}}}, \qquad \tilde{a}_{g,i,1,j} = \alpha_{g,i,1,j}\zeta_{1,j}, \qquad R_{g,i}^{\text{reg}} = \mathcal{R}((\mathbf{M}_g^{\text{reg}})_{iK}).$$

If the loading factors were stationary (time-independent), we would have for any time t

$$X_t^{(q)} = \boldsymbol{a}_{g,i}^{\text{reg}} \cdot \boldsymbol{Z}_t + \sqrt{1 - \boldsymbol{a}_{g,i}^{\text{reg}} \cdot \mathbf{C}\boldsymbol{a}_{g,i}^{\text{reg}}}\,\varepsilon_t^{(q)},$$

and the first approach would be valid. However, the micro-correlation and macro-correlation parameters evolve in time so we need to update this representation.
The unnormalized log asset value is given by

$$\tilde{X}_t^{(q)} = \tilde{\boldsymbol{a}}_{g,i,t} \cdot \boldsymbol{Z}_t + \tilde{\sigma}_{g,i}\varepsilon_t^{(q)},$$

which is a Gaussian variable with mean zero and variance $\tilde{\boldsymbol{a}}_{g,i,t} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,t} + \tilde{\sigma}_{g,i}^2$. Here $\tilde{a}_{g,i,t,j} = \alpha_{g,i,t,j}\zeta_{t,j}$ and $\tilde{\sigma}_{g,i}$ does not depend on t because the unnormalized idiosyncratic risk is assumed to be stationary.

The normalized log asset value $X_1^{(q)}$ needs to be of variance one so that the migration matrix $\mathbf{M}_{g,1}$ is equal to $\mathbf{M}_g^{\text{reg}}$. This means that $X_1^{(q)} = \tilde{X}_1^{(q)}/\sqrt{\tilde{\boldsymbol{a}}_{g,i,1} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,1} + \tilde{\sigma}_{g,i}^2}$. Since $X_1^{(q)}$ is of the form (47), the variance $\tilde{\sigma}_{g,i}^2$ solves

$$1 - \boldsymbol{a}_{g,i}^{\text{reg}} \cdot \mathbf{C}\boldsymbol{a}_{g,i}^{\text{reg}} = \frac{\tilde{\sigma}_{g,i}^2}{\tilde{\boldsymbol{a}}_{g,i,1} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,1} + \tilde{\sigma}_{g,i}^2},$$

which gives, with the identity $R_{g,i}^{\text{reg}} = \boldsymbol{a}_{g,i}^{\text{reg}} \cdot \mathbf{C}\boldsymbol{a}_{g,i}^{\text{reg}}$,

$$\tilde{\sigma}_{g,i}^2 = \tilde{\boldsymbol{a}}_{g,i,1} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,1} \frac{1 - R_{g,i}^{\text{reg}}}{R_{g,i}^{\text{reg}}}.$$

The normalized log asset value $X_t^{(q)}$ needs to be of variance one so that the migration matrix $\mathbf{M}_{g,t}$ is equal to $\mathbf{M}_g^{\text{reg}}$. This means that $X_t^{(q)} = \tilde{X}_t^{(q)}/\sqrt{\tilde{\boldsymbol{a}}_{g,i,t} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,t} + \tilde{\sigma}_{g,i}^2}$. Since $X_t^{(q)}$ is of the form (1), we find that the factor loadings $a_{g,i,t,j}$ are of the form

$$a_{g,i,t,j} = \frac{\tilde{a}_{g,i,t,j}}{\sqrt{\tilde{\boldsymbol{a}}_{g,i,t} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,t} + \tilde{\sigma}_{g,i}^2}},$$

which gives (45), and the correlation is of the form

$$R_{g,i,t} = \frac{\tilde{\boldsymbol{a}}_{g,i,t} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,t}}{\tilde{\boldsymbol{a}}_{g,i,t} \cdot \mathbf{C}\tilde{\boldsymbol{a}}_{g,i,t} + \tilde{\sigma}_{g,i}^2}.$$

which gives (44). □

Remark. Note that, in this approach, the correlation $R_{g,i,t}$ is different from (typically, larger than) $R_{g,i}^{\text{reg}}$, which means that the exposition to the systematic risk factors is different from (typically larger than) the exposition defined by the regulator. As the migration matrices are assumed to be constant and given by $\mathbf{M}_g^{\text{reg}}$, this means that the exposition to the idiosyncratic risks $\sqrt{1 - R_{g,i,t}^2}$ is different (typically, smaller than) the exposition defined by the regulator. To sum-up, if the climatic risk intensities increase, then the idiosyncratic risk decays in order to maintain the same unconditional migration matrices. This may be problematic.

Under these hypotheses, the expected loss is given by (6):

$$L^{\mathrm{e}} = \sum_{t=1}^{T} L_t^{\mathrm{e}},$$
$$L_1^{\mathrm{e}} = \sum_{g=1}^{G} \sum_{i=1}^{K-1} (\mathbf{M}_g^{\mathrm{reg}})_{iK} \mathrm{LGD}_{g,i}^{\mathrm{reg}} \mathrm{EAD}_{g,i,1},$$
$$L_t^{\mathrm{e}} = \sum_{g=1}^{G} \sum_{i,j=1}^{K-1} \big((\mathbf{M}_g^{\mathrm{reg}})^{t-1}\big)_{ij} \big(\mathbf{M}_g^{\mathrm{reg}}\big)_{jK} \mathrm{LGD}_{g,j}^{\mathrm{reg}} \mathrm{EAD}_{g,i,t},$$

for $t \geq 2$, and the conditional loss given a trajectory $\mathbf{Z} = (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_T)$ of the systematic risk factors is given by (20):

$$L(\mathbf{Z}) = \sum_{t=1}^{T} L_t(\mathbf{Z}),$$
$$L_1(\mathbf{Z}) = \sum_{g=1}^{G} \sum_{i=1}^{K-1} (\mathbf{M}_g^{\mathrm{reg}}(\boldsymbol{Z}_1))_{iK} \mathrm{LGD}_{g,i}^{\mathrm{reg}} \mathrm{EAD}_{g,j,1},$$
$$L_t(\mathbf{Z}) = \sum_{g=1}^{G} \sum_{i,j=1}^{K-1} \big(\mathbf{M}_g^{\mathrm{reg}}(\boldsymbol{Z}_1) \cdots \mathbf{M}_g^{\mathrm{reg}}(\boldsymbol{Z}_{t-1})\big)_{ij} \big(\mathbf{M}_g^{\mathrm{reg}}(\boldsymbol{Z}_t)\big)_{jK} \mathrm{LGD}_{g,j}^{\mathrm{reg}} \mathrm{EAD}_{g,i,t},$$

for $t \geq 2$, where the conditional migration matrices are given by (28). The conditional loss in stressed conditions $L_{\mathrm{stress}}^{1-\alpha}$ is the $1-\alpha$-quantile of $L(\mathbf{Z})$ when the $\boldsymbol{Z}_t$ are independent and identically distributed with the multivariate normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{C})$ with $\mathbf{C}$ given by (36).

6.2.3 *Approach 3 for the correlation model and loading factors.*

We consider here that:
- the time unit is one year,
- at time 1 the migration matrix $\mathbf{M}_{g,1}$ is equal to $\mathbf{M}_g^{\mathrm{reg}}$ and the correlation $R_{g,i,1}$ is determined by the regulator's formula,
- the migration matrices and the regulator's formula for the correlation are updated at time $t \geq 2$ because, contrary to the economic and idiosyncratic risks, which are stationary, the physical and transition risks evolve in time.
- the factor loadings $a_{g,i,t,j}$ are proportional to the product of the macro-correlation and micro-correlation adjustment parameters.

As a result, at time 1, the formulas are reduced to the formulas of the first approach:

$$\mathbf{M}_{g,1} = \mathbf{M}_g^{\mathrm{reg}}, \tag{48}$$

$$R_{g,i,1} = R_{g,i}^{\mathrm{reg}}, \qquad R_{g,i}^{\mathrm{reg}} = \mathcal{R}((\mathbf{M}_g^{\mathrm{reg}})_{iK}), \tag{49}$$

$$a_{g,i,1,j} = a_{g,i,j}^{\mathrm{reg}}, \qquad a_{g,i,j}^{\mathrm{reg}} = \sqrt{R_{g,i}^{\mathrm{reg}}} \frac{\tilde{a}_{g,i,1,j}}{\sqrt{\tilde{\boldsymbol{a}}_{g,i,1} \cdot \mathbf{C} \tilde{\boldsymbol{a}}_{g,i,1}}}, \tag{50}$$

with $\mathcal{R}$ defined by (38) and

$$\tilde{a}_{g,i,t,j} = \alpha_{g,i,t,j} \zeta_{t,j}. \tag{51}$$

At time $t \geq 1$, we have

$$(\mathbf{M}_{g,t})_{ij} = \begin{cases} 1 - \Phi(z_{g,t,i2}) & \text{if } j = 1, \\ \Phi(z_{g,t,ij}) - \Phi(z_{g,t,ij+1}) & \text{if } 2 \leq j \leq K-1, \\ \Phi(z_{g,t,iK}) & \text{if } j = K, \end{cases} \tag{52}$$

with

$$z_{g,t,ij} = \frac{z_{g,ij}^{\mathrm{reg}}}{\sqrt{1 + \boldsymbol{c}_{g,i,t} \cdot \mathbf{C} \boldsymbol{c}_{g,i,t} - \boldsymbol{a}_{g,i}^{\mathrm{reg}} \cdot \mathbf{C} \boldsymbol{a}_{g,i}^{\mathrm{reg}}}}, \tag{53}$$

$$z_{g,ij}^{\mathrm{reg}} = \Phi^{-1}\Big(\sum_{j'=j}^{K} (\mathbf{M}_g^{\mathrm{reg}})_{ij'} \Big), \tag{54}$$

$$c_{g,i,t,j} = a_{g,i,j}^{\mathrm{reg}} \frac{\tilde{a}_{g,i,t,j}}{\tilde{a}_{g,i,1,j}}, \tag{55}$$

and we have

$$R_{g,i,t} = \frac{\boldsymbol{c}_{g,i,t} \cdot \mathbf{C} \boldsymbol{c}_{g,i,t}}{1 + \boldsymbol{c}_{g,i,t} \cdot \mathbf{C} \boldsymbol{c}_{g,i,t} - \boldsymbol{a}_{g,i}^{\mathrm{reg}} \cdot \mathbf{C} \boldsymbol{a}_{g,i}^{\mathrm{reg}}}, \tag{56}$$

$$a_{g,i,t,j} = \frac{c_{g,i,t,j}}{\sqrt{1 + \boldsymbol{c}_{g,i,t} \cdot \mathbf{C} \boldsymbol{c}_{g,i,t} - \boldsymbol{a}_{g,i}^{\mathrm{reg}} \cdot \mathbf{C} \boldsymbol{a}_{g,i}^{\mathrm{reg}}}}. \tag{57}$$

Note that the formulas (57) and (45) for the loading factors coincide, and the formulas (56) and (44) for the correlations coincide. The difference between the approaches 2 and 3 is that the migration matrices are constant in the approach 2 and they evolve in the approach 3.

Proof. At time 1 (see the first approach) the normalized log asset value is given by

$$X_1^{(q)} = \boldsymbol{a}_{g,i}^{\mathrm{reg}} \cdot \boldsymbol{Z}_1 + \sqrt{1 - \boldsymbol{a}_{g,i}^{\mathrm{reg}} \cdot \mathbf{C} \boldsymbol{a}_{g,i}^{\mathrm{reg}}} \varepsilon_1^{(q)}.$$

If the loading factors were stationary (time-independent), we would have for any time t

$$X_t^{(q)} = \boldsymbol{a}_{g,i}^{\text{reg}} \cdot \boldsymbol{Z}_t + \sqrt{1 - \boldsymbol{a}_{g,i}^{\text{reg}} \cdot \mathbf{C}\boldsymbol{a}_{g,i}^{\text{reg}}}\,\varepsilon_t^{(q)},$$

and the first approach would be valid. However, the idiosyncratic risk is stationary but the micro-correlation and macro-correlation parameters evolve in time. This means that, using (55), we have in fact

$$\overline{X}_t^{(q)} = \boldsymbol{c}_{g,i,t} \cdot \boldsymbol{Z}_t + \sqrt{1 - \boldsymbol{a}_{g,i}^{\text{reg}} \cdot \mathbf{C}\boldsymbol{a}_{g,i}^{\text{reg}}}\,\varepsilon_t^{(q)},$$

which is a Gaussian variable with mean zero and variance $1 - \boldsymbol{a}_{g,i}^{\text{reg}} \cdot \mathbf{C}\boldsymbol{a}_{g,i}^{\text{reg}} + \boldsymbol{c}_{g,i,t} \cdot \mathbf{C}\boldsymbol{c}_{g,i,t}$. As a consequence, the probabilities of rating change are

$$(\mathbf{M}_{g,t})_{ij} = \mathbb{P}\big(\overline{X}_t^{(q)} \in [z_{g,ij+1}^{\text{reg}}, z_{g,ij}^{\text{reg}}]\big),$$

where the $z_{g,ij}^{\text{reg}}$'s are the threshold values associated to the given unconditional migration matrix $\mathbf{M}_g^{\text{reg}}$. This gives (52). Furthermore, after normalization, the log asset value $X_t^{(q)} = \overline{X}_t^{(q)} / \sqrt{1 - \boldsymbol{a}_{g,i}^{\text{reg}} \cdot \mathbf{C}\boldsymbol{a}_{g,i}^{\text{reg}} + \boldsymbol{c}_{g,i,t} \cdot \mathbf{C}\boldsymbol{c}_{g,i,t}}$ has now the form

$$X_t^{(q)} = \boldsymbol{a}_{g,i,t} \cdot \boldsymbol{Z}_t + \sqrt{1 - \boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{a}_{g,i,t}}\,\varepsilon_t^{(q)},$$

with $\boldsymbol{a}_{g,i,t}$ given by (57), which also gives (56). □

Remark. Note that, in this approach, if the climatic risk intensities increase, then the idiosyncratic risk and the economic risk stay constant, so that the overall risk increases and the unconditional migration matrices change.

Under these hypotheses, the expected loss is given by (6):

$$L^{\text{e}} = \sum_{t=1}^{T} L_t^{\text{e}}, \tag{58}$$

$$L_1^{\text{e}} = \sum_{g=1}^{G} \sum_{i=1}^{K-1} (\mathbf{M}_{g,1})_{iK} \text{LGD}_{g,i}^{\text{reg}} \text{EAD}_{g,i,1}, \tag{59}$$

$$L_t^{\text{e}} = \sum_{g=1}^{G} \sum_{i,j=1}^{K-1} \big(\mathbf{M}_{g,1} \cdots \mathbf{M}_{g,t-1}\big)_{ij} \big(\mathbf{M}_{g,t}\big)_{jK} \text{LGD}_{g,j}^{\text{reg}} \text{EAD}_{g,i,t}, \tag{60}$$

for $t \geq 2$, and the conditional loss given a trajectory $\mathbf{Z} = (\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_T)$ of the systematic risk factors is given by (20):

$$L(\mathbf{Z}) = \sum_{t=1}^{T} L_t(\mathbf{Z}), \tag{61}$$

$$L_1(\mathbf{Z}) = \sum_{g=1}^{G} \sum_{i=1}^{K-1} (\mathbf{M}_{g,1}(\boldsymbol{Z}_1))_{iK} \mathrm{LGD}_{g,i}^{\mathrm{reg}} \mathrm{EAD}_{g,j,1}, \tag{62}$$

$$L_t(\mathbf{Z}) = \sum_{g=1}^{G} \sum_{i,j=1}^{K-1} \big(\mathbf{M}_{g,1}(\boldsymbol{Z}_1) \cdots \mathbf{M}_{g,t-1}(\boldsymbol{Z}_{t-1})\big)_{ij} \big(\mathbf{M}_{g,t}(\boldsymbol{Z}_t)\big)_{jK} \mathrm{LGD}_{g,j}^{\mathrm{reg}} \mathrm{EAD}_{g,i,t}, \tag{63}$$

for $t \geq 2$, where the conditional migration matrices are given by (28). The conditional loss in stressed conditions $L_{\mathrm{stress}}^{1-\alpha}$ is the $1-\alpha$-quantile of $L(\mathbf{Z})$ when the $\boldsymbol{Z}_t$ are independent and identically distributed with the multivariate normal distribution $\mathcal{N}(\mathbf{0}, \mathbf{C})$ (with $\mathbf{C}$ given by (36) for instance).

7 Perspectives: Towards reverse stress test ?

The strategy developed in Section 4 makes it possible to estimate the conditional loss under stressed conditions $L_{\mathrm{stress}}^{1-\alpha}$. We may want to determine which systematic risk is the most important or which types of trajectories are the most likely to lead to a loss that exceeds $L_{\mathrm{stress}}^{1-\alpha}$. For this, we can look for the conditional distribution of the process $\mathbf{Z}$ given $L(\mathbf{Z}) \geq L_{\mathrm{stress}}^{1-\alpha}$. We may in particular want to determine $\mathbb{E}[\boldsymbol{Z}_t | L(\mathbf{Z}) \geq L_{\mathrm{stress}}^{1-\alpha}]$ for $t = 1, \ldots, T$. This could be estimated by a straightforward use of the Monte Carlo sample generated for the estimation of the quantile $L_{\mathrm{stress}}^{1-\alpha}$.

A Annex: The special recovery model (2)

The model (2) allows for flexibility, easy manipulation, and (relatively) easy calibration. It uses the cumulative Gaussian distribution function Φ for easy calculations, by the two following Gaussian formulas:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Phi(ax+b) \exp\big(-\frac{x^2}{2}\big) dx = \Phi\Big(\frac{b}{\sqrt{1+a^2}}\Big), \tag{64}$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c} \Phi(ax+b) \exp\big(-\frac{x^2}{2}\big) dx = \Phi_2\Big(\frac{b}{\sqrt{1+a^2}}, c; -\frac{a}{\sqrt{1+a^2}}\Big), \tag{65}$$

where Φ is the cdf of the standard Gaussian distribution and $\Phi_2(\cdot,\cdot;\rho)$ is the bivariate cumulative Gaussian distribution with correlation ρ.

The calibration of the model can be performed using the following results.

A.1 The rank correlation

The Kendall's Tau (rank correlation) of a pair of random variables (X, Y) is defined by

$$\tau(X,Y) = \mathbb{P}\big((X-\tilde{X})(Y-\tilde{Y}) > 0\big) - \mathbb{P}\big((X-\tilde{X})(Y-\tilde{Y}) < 0\big), \tag{66}$$

when (X, Y) and $(\tilde{X}, \tilde{Y})$ are independent and identically distributed.

From (1) and (2), we get that the Pearson (linear) correlation coefficient between X and $\Phi^{-1}(R)$ is

$$\rho(X_t^{(q)}, \Phi^{-1}(\mathrm{RR}_t^{(q)})) = \boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{b}_{g,i,t}. \tag{67}$$

Since $(X, \Phi^{-1}(R))$ is a Gaussian vector, the Kendall's Tau is related to the Pearson correlation coefficient through the Greiner's equality:

$$\tau(X_t^{(q)}, \Phi^{-1}(\mathrm{RR}_t^{(q)})) = \frac{2}{\pi}\arcsin\big(\rho(X_t^{(q)}, \Phi^{-1}(\mathrm{RR}_t^{(q)}))\big). \tag{68}$$

The Kendall's Tau is invariant under strictly increasing transform, so

$$\begin{aligned} \tau(X_t^{(q)}, \mathrm{RR}_t^{(q)}) &= \tau(X_t^{(q)}, \Phi^{-1}(\mathrm{RR}_t^{(q)})) \\ &= \frac{2}{\pi}\arcsin\big(\boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{b}_{g,i,t}\big). \end{aligned} \tag{69}$$

A.2 The Loss Given Default

It is important to note, (for calibration purposes), that the recovery rate and the default occurrence are correlated. This means that the unconditional expectation of (one minus) the recovery rate

$$\mathbb{E}\big[1 - \mathrm{RR}_t^{(q)}\big] = 1 - \Phi\Big(\frac{\mu_{g,i,t}}{\sqrt{1+\sigma_{g,i,t}^2}}\Big) \tag{70}$$

is not the expected Loss Given Default, that is observed for the borrowers who default. The expected Loss Given Default for the borrowers from group g and with rating i at time $t-1$ who default at time t is

$$\mathbb{E}\big[1 - \mathrm{RR}_t^{(q)} | X_t^{(q)} \leq z_{g,t,iK}\big] \tag{71}$$

because the event "$X_t^{(q)} \leq z_{g,t,iK}$" corresponds to default for such borrowers. The expected Loss Given Default for the borrowers from group g and with rating i at time $t-1$ who default at time t actually depends on the rating i:

$$\mathbb{E}\big[1 - \mathrm{RR}_t^{(q)} | X_t^{(q)} \leq z_{g,t,iK}\big] = 1 - \frac{1}{(\mathbf{M}_{g,t})_{iK}} \Phi_2\Big(\frac{\mu_{g,i,t}}{\sqrt{1+\sigma_{g,i,t}^2}}, z_{g,t,iK}; \frac{-\rho_{g,i,t}\sigma_{g,i,t}}{\sqrt{1+\sigma_{g,i,t}^2}}\Big), \tag{72}$$

where $\rho_{g,i,t} = \boldsymbol{a}_{g,i,t} \cdot \mathbf{C}\boldsymbol{b}_{g,i,t}$.

Proof. The distribution of $X_t^{(q)}$ is $\mathcal{N}(0,1)$. The distribution of $\Phi^{-1}(\mathrm{RR}_t^{(q)})$ is $\mathcal{N}(\mu_{g,i,t}, \sigma_{g,i,t}^2)$. The correlation coefficient between $X_t^{(q)}$ and $\Phi^{-1}(\mathrm{RR}_t^{(q)})$ is $\rho_{g,t}$. The vector $(X_t^{(q)}, \Phi^{-1}(\mathrm{RR}_t^{(q)}))$ is Gaussian, so the conditional distribution of $\Phi^{-1}(\mathrm{RR}_t^{(q)})$ given $X_t^{(q)} = x$ is $\mathcal{N}(\mu_{g,i,t} + \rho_{g,i,t}\sigma_{g,i,t}x, \sigma_{g,i,t}^2(1-\rho_{g,i,t}^2))$ and we get

$$\begin{aligned}\mathbb{E}[\mathrm{RR}_t^{(q)} | X_t^{(q)} = x] &= \frac{1}{\sqrt{2\pi\sigma_{g,i,t}^2(1-\rho_{g,i,t}^2)}} \int_{-\infty}^{\infty} \Phi(r) \exp\Big(-\frac{(r-\mu_{g,i,t}-\rho_{g,i,t}\sigma_{g,i,t}x)^2}{2\sigma_{g,i,t}^2(1-\rho_{g,i,t}^2)}\Big) dr \\ &= \Phi\Big(\frac{\mu_{g,i,t}+\rho_{g,i,t}\sigma_{g,i,t}x}{\sqrt{1+\sigma_{g,i,t}^2(1-\rho_{g,i,t}^2)}}\Big)\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}\big[\mathrm{RR}_t^{(q)} | X_t^{(q)} \leq z_{g,t,iK}\big] &= \frac{\mathbb{E}\big[\mathrm{RR}_t^{(q)} \mathbf{1}_{X_t^{(q)} \leq z_{g,t,iK}}\big]}{\mathbb{P}(X_t^{(q)} \leq z_{g,t,iK})} \\ &= \frac{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_{g,t,iK}} \Phi\Big(\frac{\mu_{g,i,t}+\rho_{g,i,t}\sigma_{g,i,t}x}{\sqrt{1+\sigma_{g,i,t}^2(1-\rho_{g,i,t}^2)}}\Big) \exp\big(-\frac{x^2}{2}\big) dx}{(\mathbf{M}_{g,t})_{ik}},\end{aligned}$$

which gives (72). □

Similarly, we have

$$\mathbb{E}\big[\Phi^{-1}(\mathrm{RR}_t^{(q)})|X_t^{(q)} \le z_{g,t,iK}\big] = \mu_{g,i,t} - \rho_{g,i,t}\sigma_{g,i,t}\frac{\exp(-z_{g,t,iK}^2/2)}{\sqrt{2\pi}(\mathbf{M}_{g,t})_{iK}}, \tag{73}$$

$$\begin{aligned}\mathbb{E}\big[\Phi^{-1}(\mathrm{RR}_t^{(q)})^2|X_t^{(q)} \le z_{g,t,iK}\big] =& \mu_{g,i,t}^2 - 2\rho_{g,i,t}\sigma_{g,i,t}\mu_{g,i,t}\frac{\exp(-z_{g,t,iK}^2/2)}{\sqrt{2\pi}(\mathbf{M}_{g,t})_{iK}} \\ &- \sigma_{g,i,t}^2\rho_{g,i,t}^2 z_{g,t,iK}\frac{\exp(-z_{g,t,iK}^2/2)}{\sqrt{2\pi}(\mathbf{M}_{g,t})_{iK}} + \sigma_{g,i,t}^2.\end{aligned} \tag{74}$$

Proof. These formulas follow from the fact that the conditional distribution of $\Phi^{-1}(\mathrm{RR}_t^{(q)})$ given $X_t^{(q)} = x$ is $\mathcal{N}(\mu_{g,i,t} + \rho_{g,i,t}\sigma_{g,i,t}x, \sigma_{g,i,t}^2(1-\rho_{g,i,t}^2))$, so that

$$\begin{aligned}\mathbb{E}\big[\Phi^{-1}(\mathrm{RR}_t^{(q)})|X_t^{(q)} = x\big] =& \mu_{g,i,t} + \rho_{g,i,t}\sigma_{g,i,t}x, \\ \mathbb{E}\big[\Phi^{-1}(\mathrm{RR}_t^{(q)})^2|X_t^{(q)} \le z_{g,t,iK}\big] =& \mu_{g,i,t}^2 + 2\rho_{g,i,t}\sigma_{g,i,t}\mu_{g,i,t}x \\ &+ \rho_{g,i,t}^2\sigma_{g,i,t}^2x^2 + \sigma_{g,i,t}^2(1-\rho_{g,i,t}^2).\end{aligned}$$

We get the desired results by using the following Gaussian identities:

$$\begin{aligned}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z_{g,t,iK}} \exp\big(-\frac{x^2}{2}\big)dx &= (\mathbf{M}_{g,t})_{iK}, \\ \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z_{g,t,iK}} x\exp\big(-\frac{x^2}{2}\big)dx &= -\frac{1}{\sqrt{2\pi}}\exp\big(-\frac{z_{g,t,iK}^2}{2}\big), \\ \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z_{g,t,iK}} x^2\exp\big(-\frac{x^2}{2}\big)dx &= -\frac{z_{g,t,iK}}{\sqrt{2\pi}}\exp\big(-\frac{z_{g,t,iK}^2}{2}\big) + (\mathbf{M}_{g,t})_{iK}.\end{aligned}$$

□

The formulas (69), (72), (73), and (74) can be used to calibrate the parameters $\mu_{g,i,t}$, $\sigma_{g,i,t}$, and $\boldsymbol{b}_{g,i,t}$ (or $\lambda_{g,i,t}$ if we use the simplified model $\boldsymbol{b}_{g,i,t} = \lambda_{g,i,t}\boldsymbol{a}_{g,i,t}$) of the recovery model.

The conditional Loss Given Default for the borrowers from group g and with rating i at time $t-1$ who default at time t given $\boldsymbol{Z}_t$ is simple, because the correlation between the recovery rate and the default occurrence happens only through the systematic risk factors, so $X_t^{(q)}$ and $\mathrm{RR}_t^{(q)}$ are independent given $\boldsymbol{Z}_t$. The conditional loss given default for the borrowers from group g

and with rating i who default at time t given $\boldsymbol{Z}_t$ is given by (29):

$$\begin{aligned}\mathrm{LGD}_{g,i,t}(\boldsymbol{Z}_t) &:= \mathbb{E}\big[1 - \mathrm{RR}_t^{(q)} | X_t^{(q)} \leq z_{g,t,iK}, \boldsymbol{Z}_t\big] \\ &= \mathbb{E}\big[1 - \mathrm{RR}_t^{(q)} | \boldsymbol{Z}_t\big] \\ &= 1 - \Phi\Big(\frac{\mu_{g,i,t} + \sigma_{g,i,t}\boldsymbol{b}_{g,i,t} \cdot \boldsymbol{Z}_t}{\sqrt{1 + \sigma_{g,i,t}^2(1 - \boldsymbol{b}_{g,i,t} \cdot \mathbf{C}\boldsymbol{b}_{g,i,t})}}\Big). \end{aligned} \tag{75}$$

References

[1] L. Andersen and J. Sidenius, Extensions to the Gaussian Copula: Random recovery and random factor loadings, Journal of Credit Risk **1** (2004), 29–70. 4

[2] M. Gordy, A risk-factor model foundation for ratings-based bank capital rules, Journal of Financial Intermediation **12** (2003), 199–232. 10

[3] Green RWA, How banks can save the planet, Part 1 - July 2020. 2

[4] J. C. Hull and A. D. White, Valuation of a CDO and an n-th to default CDS without Monte Carlo simulation, The Journal of Derivatives **12** (2004), 8–23. 3

[5] R. C. Merton, On the pricing of corporate debt: the risk structure of interest rates, Journal of Finance **29** (1974), 449–470. 3

[6] M. Pykhtin, Portfolio credit risk multi-factor adjustment, Risk **17** (2004), 85–90. 2

[7] M. Pykhtin and A. Dev, Credit risk in asset securitisations: an analytical model, Risk **15** (2002), 16–20. 10

[8] T. Roncalli, Handbook of Financial Risk Management, CRC Press, Boca Raton, 2020. 2

[9] O. Vasicek, Loan portfolio value, Risk **15** (2002), 160–162. 2, 3, 10