FINANCIAL ENGINEERING
Final Project Report

Model risk on Regulatory Capital

Callini Matteo - Delera Giacomo

A.Y. 2023/2024

POLITECNICO
MILANO 1863

Contents

1 Introduction

The set of international banking regulations mandates that banks maintain a minimum capital reserve as a safeguard against potential future losses, which are mainly influenced by two factors: stochastic variables risk and the uncertainty inherent in the chosen model itself.

In this project, applying the Vasicek model to a homogeneous portfolio, we focused on the measure of a portfolio credit risk. Additionally, we examined the risks associated with errors in model parameters and assessed the impact of each parameter's uncertainty on the capital requirement.
In particular, we evaluated the difference between a simplified "naive" model, which is widely used, and some stressed models where we did not considered fixed parameters, but we tried to simulate them in different ways, according to the Montecarlo approach, in order to detect the impact on the Regulatory capital.
We employed a confidence level $\alpha = 99.9\%$ as established by the Basel Committee, but we also stressed the results with $\alpha = 99\%$ in every part of the project.

In the first section, we conducted some statistical tests that verified the assumptions needed for the following parts, particularly those regarding the gaussianity assumptions. However, in the final part of the report, we also used different distributional assumptions regarding the parameters in order to detect the differences.

We considered both an asymptotic homogeneous portfolio (LHP) and a homogeneous portfolio (HP) with a different finite number of obligors to verify the differences in the case of a small number of obligors (N = 50) and the possible relations in the case of a large number (e.g., N = 1000).

In addition, we compared the Regulatory capitals derived from the Standard model to those obtained through the IRB model, which is the one employed in this discussion.

In this project we mainly followed the structure proposed in the paper "The measure of model risk in credit capital requirements"[1].

Total computation time for the Matlab code: 6 minutes and 15 seconds.

1.1 Dataset

The dataset, sourced from Moody's public data, spans the years 1983-2019, so it contains 37 observations, and includes the one-year Default Rate (DR) for Speculative-grade issuers (B-rated corporate bonds) and All Rated issuers (BBB-rated corporate bonds): both parameters (PD and LGD) represent forecasts over a one-year time horizon.
Moreover, it contains the annual Recovery rates for senior unsecured corporate bond.

We considered the correlation between assets as a deterministic function of Probability of Default (PD) as specified by the Basel Committee.

2 Statistical analysis and Calibration

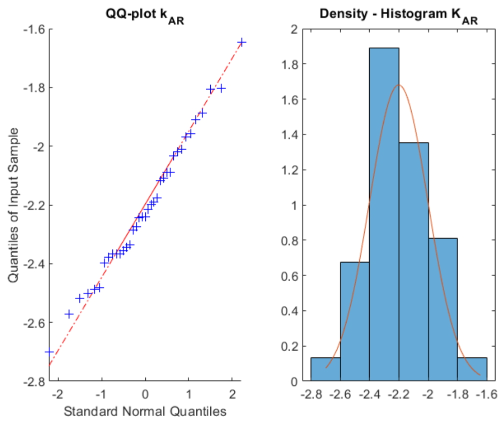
In the first part, we statistically tested the distributional assumptions on the Probabilities of Default (PD) and on the Loss Given Default (LGD): instead of PD, we considered $k = \phi^{-1}(PD)$, where ϕ is the cumulative distribution function of a gaussian.
In particular, we had to verify if they followed a gaussian distribution, in both the univariate and bivariate cases, using the Shapiro-Wilk test for the former and the Royston test for the latter.

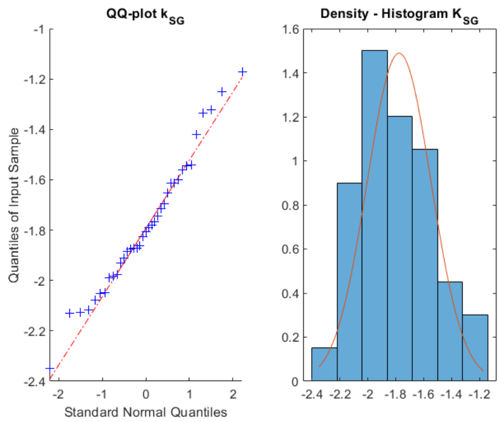
2.1 Shapiro - Wilk test

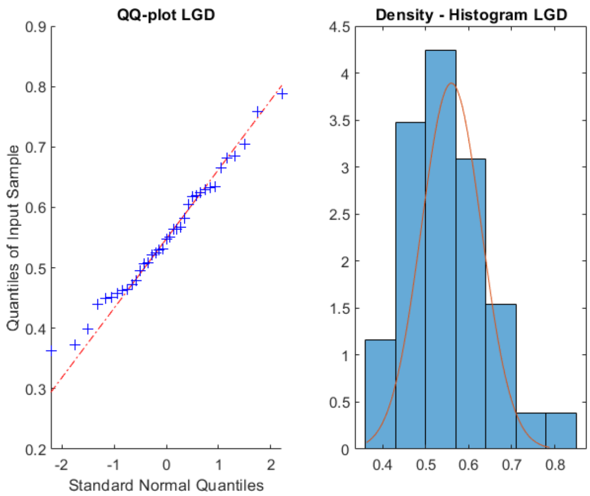
The Shapiro - Wilk test verifies the null hypothesis that a sample $x_1, ..., x_n$ follows a normal distribution.
In order to compute the test, we used the Matlab function *swtest.m* which returns the p-Value of the test, the statistical test W and H = 0 if the null hypothesis is verified or 1 if it is rejected.
We obtained the following results:

	p-Value	W
k_{SG}	0.7056	0.9793
k_{AR}	0.9413	0.9873
LGD	0.8400	0.9833

Since the significance level is 0.001, we did not reject the null hypothesis.
Moreover, we noticed that all the p-Values are greater than 0.5, so we have no reason to reject the test hypothesis and we can consider the gaussian marginal distribution for the parameters.







From the plots we observed that all the parameters are well approximated by a gaussian as we can see from the histograms.
Moreover, from the QQ - plots we noticed that all the parameters follow a gaussian distribution

since they approximately follow the straight line: in particular, the LGD and k_{SG} distributions have slightly fatter tails than the normal case since the points at the end of the QQ plot deviate from the straight line; this could potentially account for their smaller p-value, which nonetheless remains high.

2.2 Royston test

The Royston test verifies the null hyphotesis that m variates with sample size n follow a multivariate gaussian distribution.
In order to compute the test we used the Matlab function *Roystest.m* which returns the p-Value of the test and H = 0 if the null hypothesis is verified and H = 1 if it is rejected.
First, considering the Shapiro-Wilk statistical test W, the function computes a transformation z_i for each variate i=1, ..., m, which, according to Royston (1982)[2], can be approximated as a standard gaussian in function of the parameters λ, μ and σ, the last two represent the mean and the standard deviation of the transformation, which are function of the sample size n.
Then, in the function *Roystest.m*, it is defined, for each variate, the new parameter $k_i = (\phi^{-1}(\frac{1}{2}\phi(-z_i)))^2$, where ϕ is the cumulative density function of a standard gaussian.
In Royston (1983)[3], it is defined the associated statistical test:

$$H = \frac{e}{m}\sum_{i=1}^{m} k_i \sim \chi^2(e)$$

where $e = \frac{m}{1+(m-1)\bar{c}}$, $\bar{c} = \sum\sum_{i,j=1}^{m} \frac{c_{ij}}{m^2-m}$ and c_{ij} are obtained considering a transformation of the sample correlation matrix described in Royston (1983)[3].
The p-Value is given by 1 - *chi2cdf*(H,e), where *chi2cdf* is the cumulative distribution function of a χ^2 with e degrees of freedom.
We obtained the following results:

	p-Value
$k_{SG} - LGD$	0.9141
$k_{AR} - LGD$	0.9756

As before, we obtained p-Values way larger than 0.1 and clearly greater than $\alpha = 0.001$: so, we did not reject the null hypothesis and we could consider the couples $k_{SG} - LGD$ and $k_{AR} - LGD$ as bivariate gaussian.

2.3 Pearson correlation coefficients

We analyzed the relation between the parameters: in particular, we computed the Pearson correlation coefficient between two generic samples x and y, which is defined as:

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$$

Moreover, we computed the 95% confidence interval for the coefficient.
First, assuming a bivariate normal population with size n, we defined the transformation of the sample product moment correlation from r to z_r:

$$z_r = \frac{1}{2}ln(\frac{1+r}{1-r})$$

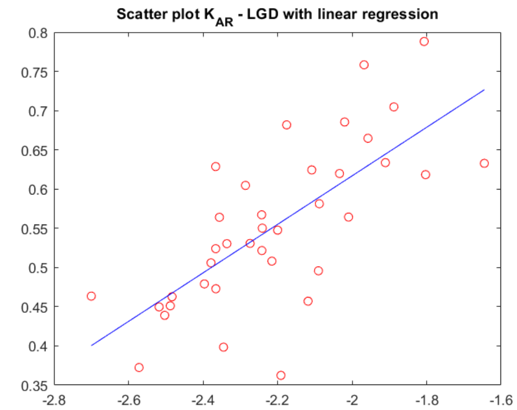
which is approximately distributed as a gaussian with variance $\frac{1}{n-3}$ (Bonett and Wright (2000)[4]).
Then we defined the confidence interval for z as $z_r \pm z_{1-\alpha/2}\sqrt{\frac{1}{n-3}}$.
So, inverting the initial formula for z_r, we obtained, respectively, the upper and the lower bound for r as:

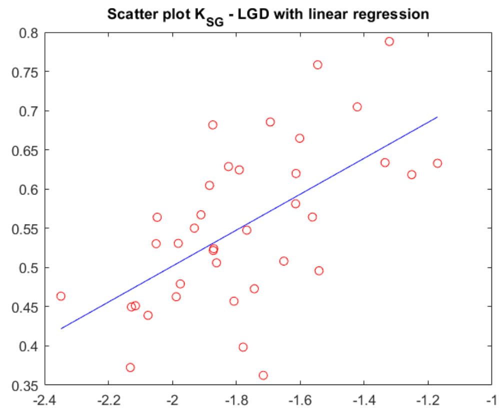
$$r_L = \frac{e^{2z_L}-1}{e^{2z_L}+1}$$

$$r_U = \frac{e^{2z_U}-1}{e^{2z_U}+1}$$

We obtained the following results:

	r	CI ($\alpha = 0.95\%$)
$k_{SG} - LGD$	0.5994	[0.3417, 0.7732]
$k_{AR} - LGD$	0.7165	[0.5112, 0.8445]





We observed a strong positive correlation between the parameters, as we can see by the high absolute values of the correlation coefficients and the wide confidence intervals.
The lowest correlation value in both the confidence intervals is 0.3417, which is still relatively significant. This relationship could be also seen from the scatter plots, where the linear regression lines have a positive slope.
We expected this outcome because the Probability of Default and the Loss Given Default tend to rise during crisis: if more obligors default, banks will recover less money, leading to greater losses. On the other hand, if banks recover less money, it may be indicative of an ongoing default trend.

2.4 Descriptive statistics

We briefly resumed the main statistical descriptive objects for each parameter.

	min	max	mean	median	std	n
LGD	0.3625	0.7881	0.5526	0.5476	0.1024	37
PD_{AR}	0.0035	0.0500	0.0159	0.0125	0.0101	37
PD_{SG}	0.0094	0.1209	0.0430	0.0354	0.0262	37

We observed that All Rated probabilities are way smaller than the Speculative grade ones, but also the standard deviation keeps this behaviour: we expected this result since the Speculative-grade corporate bonds are more likely to default since their grade (B) is below the All Rated one (BBB). The maximum values are observed in 2009, following the mortgage subprime crisis, whereas the smallest values occur in 2007, just before the global crisis.
Concerning Loss Given Defaults, we could notice that the mean exceeds 0.5, indicating that in case of default, recovering the majority of the funds becomes difficult.

2.5 k mean calibration

After the statistical tests computation described in the first subsection, we were allowed to consider k and LGD as gaussians. In order to find the mean of k we had to invert the following formula:

$$\overline{PD} = E[\phi(k)] = \int_{-\infty}^{+\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} \phi(\overline{k} + \sigma_k x) dx$$

We opted for a numerical inversion, using the Matlab command *fzero* we obtained the following results:

	$\overline{k}$
SG	-1.7779
AR	-2.2075

3 Measure of model risk in capital requirement

Once verified all the distributional assumptions, we could compute the Regulatory Capital according to the IRB approach, which is generally defined as

$$RC = Var_\alpha[L] - E[L]$$

where L is the Loss portfolio distribution and α is the confidence level defined by the regulator, in our case $\alpha = 99.9\%$.
In the Vasicek model, considering the homogeneous portfolio case, the risk factors are defined (for each obligor i) as

$$X_i = \sqrt{\rho}M + \sqrt{1-\rho}\epsilon_i$$

M and ϵ_i are i.i.d. standard gaussians: M is the market risk factor and it is common for each obligor and ϵ_i is the idiosyncratic risk factor and it represents the obligor-specific risk component.
Similarly to Section 2, we defined the default points as $k = \phi^{-1}(PD)$: given that $X_i < k_i$ represents the default condition for the obligor i, this consideration made easier to employ more suitable distributional assumptions regarding the parameters.
In this scenario, we examined a significant number of obligors, leading us to adopt the Vasicek model for the Large Homogeneous Portfolio case.
Then, we considered the conditional expectation with respect to M, PD and LGD as

$$\mathbb{E}[L|M, LGD, PD] = LGD\phi(\frac{k-\sqrt{\rho}M}{\sqrt{1-\rho}})$$

where ρ is a deterministic function of time defined by the Basel Committee and it represents the correlation between assets.

$$\rho(PD) = 0.12\frac{1-e^{-50PD}}{1-e^{-50}} + 0.24(1-\frac{1-e^{-50PD}}{1-e^{-50}})$$

This is the Regulatory correlation function for Large corporates, where 0.12 and 0.24 represent the minimum and maximum correlation: we noticed that if the probability to default is large, then the correlation becomes smaller; so, this formula considers smaller correlation in quiet periods, while it gives higher correlation in case of crisis.

3.1 Capital requirements - Naive approach

An initial approximation for Regulatory capital considers only the expected values of PD and LGD.

$$EL^{naive} = \overline{LGD} \cdot \overline{PD}$$

$$RC^{naive} = \overline{LGD}\phi(\frac{\phi^{-1}(\overline{PD}) - \sqrt{\rho(\overline{PD})}\phi^{-1}(1-\alpha)}{\sqrt{1-\rho(\overline{PD})}}) - EL^{naive}$$

The Regulatory capital is defined as the difference between the unexpected loss in the worst case scenario and the expected loss. (Tarashev 2010 [6])

Table 1: $\alpha = 0.999$

	RC^{naive}
AR	0.0866
SG	0.1224

Table 2: $\alpha = 0.99$

	RC^{naive}
AR	0.0452
SG	0.0736

This formula is clearly too simple as it solely accounts for the mean values of parameters, disregarding any "noise" they may introduce.
Additionally, it does not address scenarios where one parameter varies while the other remains constant, nor does it address situations where both parameters change simultaneously.
For this reason we are interested in the Add-on computation on the Regulatory capital, in particular considering the uncertainty of the parameters.

3.2 Regulatory capital

We computed the right Regulatory Capital according to the IRB approach taking in account different parameters stresses.
We used the Montecarlo method to simulate N = 10^7 different scenarios for the market parameter M and we simulated N possible values for k, both SG and AR cases, and LGD: in Section 2 we statistically tested the hypothesis of gaussianity for both the parameters, which were verified, so we simulated both of them according to the tests.

$$LGD \sim N(0.5526, 0.0105)$$
$$k_{SG} \sim N(-1.7779, 0.0718)$$
$$k_{AR} \sim N(-2.2075, 0.0563)$$

We underline that in the code we did not fix the seed for the market, which is randomly selected at the beginning of the code: in this way, we obtained different values every time we ran the code.
In order to compare the results with the report we fixed a random seed = 150, but in general the market seed is not fixed as we said.
On the other hand, we fixed the seed for the parameters k and LGD since they are independent from the market.
For every stress, we simulated the parameters and we considered $\mathbb{E}[L|M, LGD, PD]$, which represents the Loss distribution given the simulated parameters, and $\rho(\phi(k))$ defined at the beginning of this section for each simulation, obtaining two N-dimensional vector for the losses and the correlations.
So, we considered the Expected Loss (EL) as the mean of this vector (we are allowed to compute the mean in this way since $\mathbb{E}[E[L|M, LGD, PD]] = \mathbb{E}[L]$ by the conditional expectation tower property) and we took the α-quantile of the loss distribution in order to define the new IRB Regulatory Capital as:

$$RC = Var_\alpha[\mathbb{E}[L|M, LGD, PD]] - \mathbb{E}[L]$$

For every stress we simulated 2 x N standard gaussian matrix, defined as g, for k and LGD and then we defined $LGD = \overline{LGD} + \sigma_{LGD} \cdot g(1, :)$ and $k = \overline{k} + \sigma_k \cdot g(2, :)$.
We considered four different cases:

- LGD simulated, k fixed: we fixed $k = \phi^{-1}(\overline{PD})$ since we are considering the expected mean of the default probabilities as fixed parameter and we simulated only LGD. Consequently, the correlation between assets is fixed too, since it is function of PD.
- k simulated, LGD fixed: we fixed LGD = $\overline{LGD}$ and we simulated only k.
- k and LGD simulated independently: we simulated both of them simultaneously.
- k and LGD simulated dependently: we consider that between the two parameters there is a correlation equal to the Pearson coefficient (Section 2.3). In order to correlate the parameters, we defined A as the lower Cholesky matrix of the correlation matrix C such that $C = AA^T$. Then, we defined $x = A \cdot g$ and we considered $LGD = \overline{LGD} + \sigma_{LGD} \cdot x(1, :)$ and $k = \overline{k} + \sigma_k \cdot x(2, :)$ and we proceeded in the same way as before.

Computations were made in both AR and SG cases.
For the Regulatory capital we also computed the Confidence intervals with $\alpha = 99.9\%$: we decided to use a large confidence level because every simulation of the market parameters slightly impacts on the Regulatory capitals, even with $N = 10^7$ simulations.
As we said before, in order to make the report comparable with the code we fixed a random seed for the market in order to have the same result, but in general the market parameter is not fixed.

Expected loss confidence interval with level α

$$\mathbb{E}[L] \pm T^{-1}(\frac{1 - \alpha_{int}}{2}, N)\sqrt{\frac{\sigma^2(EL_v)}{N}}$$

where T represents the cumulative distribution function of a t-student with N degrees of liberty and EL_v represents the loss vector obtained with Montecarlo.
Before computing the Regulatory Capital confidence interval, we defined the confidence interval for a quantile.

Quantile confidence interval with level α

$$qm_\alpha \pm T^{-1}(\frac{1-\alpha_{int}}{2}, N)\sqrt{\frac{\alpha(1-\alpha)}{(exp(-qm_\alpha^2)/\sqrt{2\pi})^2}}$$

where qm_α is the α - quantile of a generic distribution, in our case $qm_\alpha = Var_\alpha[E[L|M, LGD, PD]]$, and α_{int} is the confidence level. (Statistical Interval [7])
Since the Regulatory capital is defined as the difference between two quantities with unknown variance, we defined the weighted mean of the sample variances as $S^2 = \frac{(N-1)S^2_{quantile}+(N-1)S^2_{EL}}{2(N-1)}$.
Then we used the confidence interval formula for the difference between two quantities as

$$Var_\alpha[E[L|M, LGD, PD]] - E[L] \pm T^{-1}(\frac{1-\alpha_{int}}{2}, N)\sqrt{S^2(\frac{1}{N}+\frac{1}{N})}$$

We obtained the following results:

	RC	CI
LGD	0.1336	[0.1335, 0.1337]
k	0.1582	[0.1581, 0.1583]
Independent	0.1716	[0.1715, 0.1718]
Dependent	0.2034	[0.2033, 0.2036]

Table 1: SG case

	RC	CI
LGD	0.0917	[0.0916, 0.0919]
k	0.0975	[0.0974, 0.0977]
Independent	0.1033	[0.1032, 0.1034]
Dependent	0.1203	[0.1202, 0.1204]

Table 2: AR case

	Computational time (s)
LGD	1.039007
k	1.612589
Independent	1.665366
Dependent	1.759345

Table 3: Computational time for both AR and SG (it considers also the $\alpha = 0.99$ case)

We could clearly see that they are all higher that the naive case, since they consider the noise brought by the uncertainty of the parameters, in particular the k simulation has a greater impact than the LGD one and in case of dependency the amount of money requested increases a lot. We will make more complete comments in the following part.

As we said before, this result are given by a particular simulation of the market, in general it is not fixed: if we let the market seed free, we obtained that the Regulatory capital changes only at the fourth decimal digit.
For example, we noticed that the values in the LGD simulation case for AR are contained in (0.0914, 0.0920), from an empirical point of view.
So, the confidence intervals obtained are referred to a particular simulation and they are not comparable with all the simulations, but only with the given market parameters of the simulation.
A possible solution for this issue would be increasing the number of simulation, but the computational time would increase proportionally.

3.3 Add-on

In this section, we focused on the Add-ons, which are defined as the ratio between the excess loss reserve (inclusive of the expected loss correction) and the RC in the naïve approximation:

$$add-on := \frac{(RC - RC^{naive}) + (EL - EL^{naive})}{RC^{naive}}$$

In this case RC and EL represent the new values for the Regulatory capital after the parameters stress.
Add-on quantity the increase in capital requirement with respect to the naive approach: in particular, the new Regulatory capital is defined as RC_{new} = RC(1 + Add-on).
It is a crucial value in terms of our analysis, as it allows us to observe, in addition to the aforementioned variation in capital requirements, the impact of each parameter on it. It also enables us to verify the accuracy of the results with robustness tests, by observing its variation.
We attained the following values: (we write the Add-on in percentage value, this holds also for all the other sections)

Table 3: Add-on - AR case - $\alpha = 99.9\%$

	Add-on	CI
LGD	5.92	[5.76, 6.07]
k	12.61	[12.46, 12.76]
Independent	19.27	[19.12, 19.43]
Dependent	39.65	[39.49, 39.80]

Table 4: Add-on - SG case - $\alpha = 99.9\%$

	Add-on	CI
LGD	9.17	[9.05, 9.29]
k	29.26	[29.13, 29.38]
Independent	40.25	[40.13, 40.38]
Dependent	67.43	[67.30, 67.56]

As we said before, the Add-ons are greater for the independent and dependent simulations: in particular, the maximum values of the Add-on are close to 39% in the AR case and close to 67% for the SG case; the minimum values are respectively close to 6% and 9%
These values are important benchmarks for model risk measurement and they underline the risk of the naive approach.
As we saw in the previous case with the RCs, these values depend on the market parameter seed: we noticed that the values in the Add-on case change up to the second decimal digit. In the Add-on computation we divide for the Naive Regulatory capital, which first significant digit is the first decimal digit, so if the values change, the impact is greater.
Since the Add-on represent an adding percentage of the Regulatory capital, we expect that their value is more sensitive to the changes of the Regulatory capital.
Moreover, in the dependent case there is also a difference in Expected Loss values:

Table 5: AR case

	EL
naive	0.0088
LGD	0.0088
k	0.0088
Independent	0.0088
Dependent	0.0094

Table 6: SG case

	EL
naive	0.0237
LGD	0.0237
k	0.0237
Independent	0.0237
Dependent	0.0252

For all the simulations, the expected loss is close to naive case, while, in the dependent case, the expected loss is greater: this is another reason behind the great values obtained for this particular scenario.

3.4 Robustness check

To verify the consistency of the obtained results, we conducted two different robustness tests: in the first one, we considered a lower confidence level ($\alpha = 99\%$), and in the second one, we performed a granularity test, i.e., we considered a finite number of obligors.
What we expect to observe is a significant change in the regulatory capitals (for example, in the case of $\alpha = 99\%$, we expect lower RC since we are considering a lower confidence interval), but we do not expect major variations in the add-ons, indicating that the method we are analyzing suggests an RC sensitive to the framework.
In this section we just focus on the first robustness test.
As expected, the RCs are lower than in the case where $\alpha = 99.9\%$, since we are considering a better worst case scenario. By observing the variation in the RCs (they have become about half), we could affirm that the Add-ons remain similar in the two different cases: in particular they are all a bit smaller which is related to lower RC.
We noticed only one exception since the Add-on in the AR case simulating only k is greater than the $\alpha = 99.9\%$ case: as we said before, also the naive Regulatory capital decreased; so, in this

particular situation the Add-on is greater since the impact of the decreasing naive RC is bigger than the decreasing RC simulated.
The test with $\alpha = 99\%$ yielded the following regulatory capital and Add-ons:

AR case: $\text{RC}_{naive} = 0.0452$

Table 7: RC - AR case - $\alpha = 99.0\%$

	RC	CI
LGD	0.0467	[0.0463,0.0471]
k	0.0518	[0.0514,0.0522]
Independent	0.0534	[0.0531,0.0538]
Dependent	0.0611	[0.0608,0.0615]

SG case: $\text{RC}_{naive} = 0.0736$

Table 8: RC - SG case - $\alpha = 99.0\%$

	RC	CI
LGD	0.0779	[0.0775,0.0783]
k	0.0944	[0.0941,0.0948]
Independent	0.0990	[0.0986,0.0994]
Dependent	0.1140	[0.1136,0.1144]

Table 9: Add-on - AR case - $\alpha = 99.0\%$

	Add-on	CI
LGD	3.26	[2.41, 4.10]
k	14.56	[13.72,15.41]
Independent	18.22	[17.38,19.07]
Dependent	36.69	[35.84, 37.54]

Table 10: Add-on - SG case - $\alpha = 99.0\%$

	Add-on	CI
LGD	5.87	[5.34,6.40]
k	28.34	[27.80, 28.88]
Independent	34.56	[34.02, 35.10]
Dependent	56.93	[56.38, 57.48]

We noticed that the confidence interval are a bit larger since the intervals for $\alpha = 99\%$ are larger: the function $\alpha(1-\alpha)$, which is defined in the confidence interval, is decreasing for $\alpha > 50\%$.

3.5 Granularity check

We verified the impact of granularity, so we considered a finite number of obligors in order to detect the differences with the asymptotic portfolio. We computed the value of the portfolio with $N_{obligors} = 50$ (we proceeded with the same approach described in Loffler[5]).
We proceeded analogously to the previous point with a Montecarlo approach considering $\text{N} = 3 \cdot 10^5$ simulations for each scenario and for each parameter.
Differently from the asymptotic case, we simulated the risk factors, in particular the idiosyncratic risk factors ϵ_{ij} for each simulations j and for each obligor i. Then we considered the default condition:

$$X_{ij} < k_j$$

if the condition is satisfied, then the obligor defaulted.
We underline that we are using the homogeneous assumption, so, for each scenario, each obligor has the same Loss Given Default and the same Probability to Default: in particular, they have the same Default point k.
At this point, we had a set of losses for each market scenario: we made the assumptions that in the portfolio every obligor has the same notional equal to $\frac{1}{N_{obligors}}$.So, in order to compute the percentage loss, we simply took the mean of the loss given default for each scenario.
After this step, we obtained a N - dimensional loss vector and we proceeded in the same way as the large homogeneous portfolio computing the Expected mean, the Regulatory capitals and the Add-on with their respective confidence intervals.
We considered also the $\alpha = 99\%$ case and we obtained the following results:

Table 11: 50 obligors - AR case - $\alpha = 99.9\%$

	RC	CI
LGD	0.1154	[0.1147, 0.1160]
k	0.1128	[0.1121, 0.1135]
Independent	0.1259	[0.1252, 0.1266]
Dependent	0.1440	[0.1436, 0.1445]

Table 12: 50 obligors - SG case - $\alpha = 99.9\%$

	RC	CI
LGD	0.1608	[0.1601, 0.1616]
k	0.1863	[0.1855, 0.1870]
Independent	0.1957	[0.1950, 0.1964]
Dependent	0.2294	[0.2287, 0.2301]

	Add on	CI
LGD	33.22	[32.31, 34.12]
k	30.24	[29.33, 31.14]
Independent	45.41	[44.50, 46.33]
Dependent	67.08	[66.15, 68.01]

Table 13: 50 obligors - AR case - $\alpha = 99\%$

	Add on	CI
LGD	31.44	[30.72, 32.15]
k	52.19	[51.45, 52.93]
Independent	59.90	[59.15, 60.65]
Dependent	88.64	[87.86, 89.42]]

Table 14: 50 obligors - SG case - $\alpha = 99\%$

	RC	CI
LGD	0.0604	[0.0583, 0.0626]
k	0.0686	[0.0672, 0.0699]
Independent	0.0662	[0.0641, 0.0683]
Dependent	0.0749	[0.0728, 0.0770]

	RC	CI
LGD	0.0959	[0.0938, 0.0980]
k	0.1089	[0.1075, 0.1103]
Independent	0.1145	[0.1123, 0.1167]
Dependent	0.1303	[0.1282, 0.1325]

	Add on	CI
LGD	33.70	[28.79, 38.61]
k	51.72	[46.80, 56.64]
Independent	46.43	[41.51, 51.36]
Dependent	6721	[62.26, 72.17]

	Add on	CI
LGD	30.34	[27.21, 33.48]
k	47.98	[44.81, 51.14]
Independent	55.65	[52.47, 58.83]
Dependent	79.10	[75.89, 82.32]

We noticed that the Regulatory capitals are higher than the asymptotic case: we expected this result since a smaller portfolio is considered riskier than a larger one because every obligor has a huge impact on its value.
For the same reason all the Add-on are greater than the asymptotic case since every default becomes more relevant.
In this case the confidence interval are larger than the asymptotic case since we are considering a smaller number of simulations with the same confidence level: we proceeded in this way due to computational reasons.
Moreover, we noticed that the regulatory capitals are always greater in the SG case, the same did not hold for all the Add-on: this is due to the number of simulations and to the relation with the naive RC as we said in the Add-on section.

We decided to consider also different number of obligors ([50, 100, 250, 500, 1000]) in order to verify a possible convergence with respect to the asymptotic case.

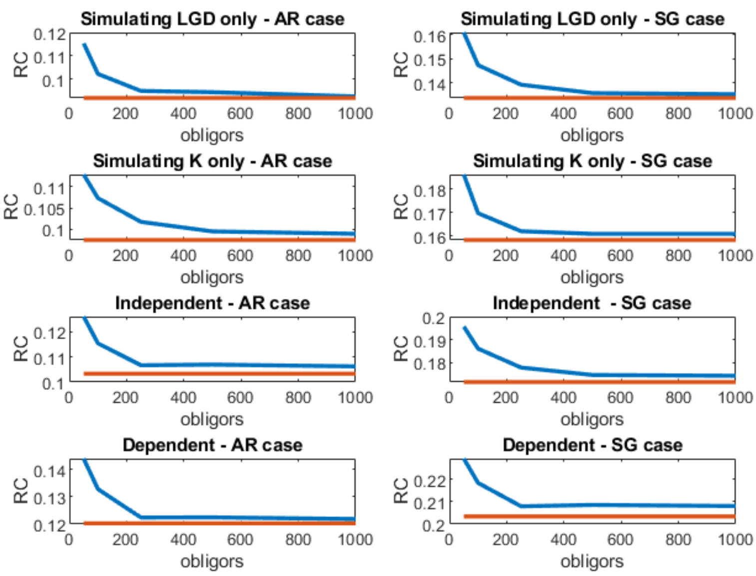


Figure 1: Regulatory capitals for $\alpha = 99.9\%$ (Red line represents the LHP RC)

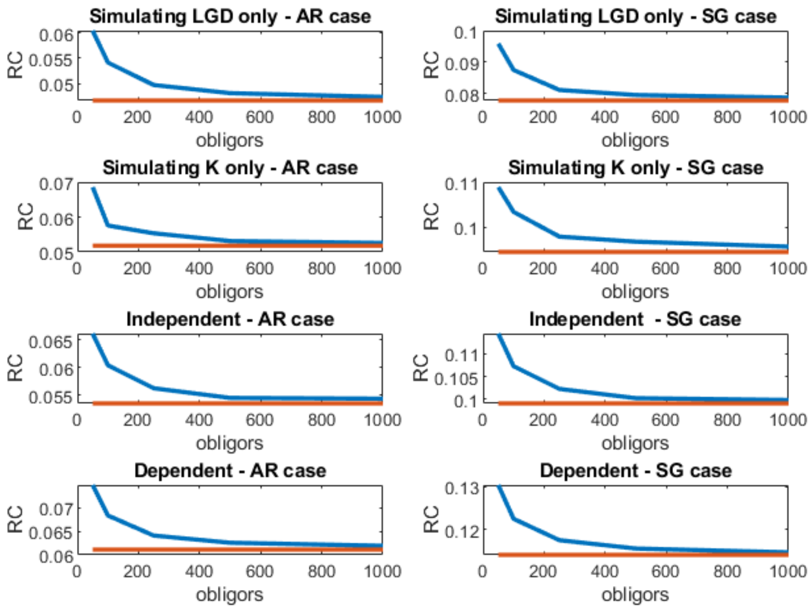


Figure 2: Regulatory capitals for $\alpha = 99\%$ (Red line represents the LHP RC)

We observed that the Regulatory capitals decrease with respect to the number of obligors: in particular, for the largest cases (500 and 1000), we noticed that the Regulatory capitals are really close to the asymptotic case, while for a small number of obligors the Regulatory capital are quite higher as we discussed in the case with 50 obligors.
As we discussed in the Robustness check part, for alpha = 99% we obtained smaller capital requirements.
Since we were computing the Montecarlo simulations with respect to $N = 3 \cdot 10^5$ scenarios, the computation is not precise if we want to compare the results with the asymptotic portfolio: indeed, we have larger confidence intervals and in some simulation of the market parameters we obtained slightly smaller Regulatory capitals for $N_{obligors} = 500$ and $N_{obligors} = 1000$ with respect to the LHP case, which is counter intuitive; but, as we said, this is due to the fewer number of simulations. We tried to increase the number of simulations, but the computation time skyrocketed since we are computing $N_{obligors} \cdot N$ risk factors.
Moreover, for alpha = 99% and $N_{obligors} = 1000$ we obtained the following values for the Add-ons.

Table 15: 1000 obligors - AR case - $\alpha = 99\%$

	Add-on	CI
LGD	4.83	[-0.02, 9.69]
k	16.28	[11.40, 21.15]
Independent	20.15	[15.27, 25.02]
Dependent	38.52	[33.61, 43.42]

Table 16: 1000 obligors - SG case - $\alpha = 99\%$

	Add-on	CI
LGD	7.13	[4.05, 10.21]
k	29.95	[26.83, 33.07]
Independent	35.60	[32.47, 38.73]
Dependent	57.73	[54.56, 60.90]

We reported these particular results since the Add-on are close to the LHP case, but the confidence intervals are way larger: in the All Rated case, in the LGD simulation, the left bound of the interval is even negative. We are considering the same confidence level for the interval, but we are using a smaller number of simulations.
So, it's better to use the asymptotic model for a large number of obligors, more than 500 in our case, since there are no relevant differences between the Regulatory capitals and the computation in the second case is lighter and even more precise, if we look at the confidence intervals.
We add some results regarding the RC in order to compare them with the asymptotic case.

Table 17: 1000 obligors - AR case - $\alpha = 99.9\%$

	RC	CI
LGD	0.0924	[0.0917, 0.0931]
k	0.0990	[0.0983, 0.0997]
Independent	0.1062	[0.1055, 0.1069]
Dependent	0.1219	[0.1212, 0.1226]

Table 18: 1000 obligors - AR case - $\alpha = 99\%$

	RC	CI
LGD	0.0474	[0.0453, 0.0495]
k	0.0526	[0.0504, 0.0545]
Independent	0.0543	[0.0522, 0.0564]
Dependent	0.0619	[0.0598, 0.0641]

Table 19: RC LHP - AR case - $\alpha = 99.9\%$

	RC	CI
LGD	0.0917	[0.0916, 0.0919]
k	0.0975	[0.0974, 0.0977]
Independent	0.1033	[0.1032, 0.1034]
Dependent	0.1203	[0.1202, 0.1204]

Table 20: RC LHP - AR case - $\alpha = 99.0\%$

	RC	CI
LGD	0.0467	[0.0463,0.0471]
k	0.0518	[0.0514,0.0522]
Independent	0.0534	[0.0531,0.0538]
Dependent	0.0611	[0.0608,0.0615]

$N_{obligors}$	Computational time (s)
50	0.547669
100	0.979805
250	2.263439
500	4.153096
1000	8.202226

Table 3: Computational time for both AR and SG (it considers also $\alpha = 0.99$ case)

3.6 Standard model

The Standard model considers the RC as a weighted sum of the asset expositions.

$$RC_{standard} = 8\% \cdot \sum_{i=1}^{n} w_i \cdot A_i$$

The Standard approach is formally accepted by the national regulators since it is simpler and more comprehensive: this approach is widely used among small banks.
We used the weights defined in Basel II.
Since we are considering an homogeneous portfolio with a large number of obligors n and we made the assumptions that all the notionals are equal to $\frac{1}{n}$, the $RC_{standard}$ is equal to 8% $\cdot$w$_i$.
In our case $w_i = 150\%$ for the Speculative grade corporate bonds (B rated) and $w_i = 100\%$ for the All Grade corporate bonds (BBB rated).
We obtained the following results:

	$RC_{standard}$
AR	0.0800
SG	0.1200

We noticed that these values are really close to the naive approach RC with $\alpha = 99.9\%$ and they are higher than the $\alpha = 99\%$ case: this could explain why the regulators decided to use this particular confidence level.
Moreover, this is a possible reason behind the widely use of the naive approach for the Regulatory capital IRB model, since this model is a slightly better approximation than the Standard one because it is based on a risk factor model and, in general, it can detect the differences between the different assets: the Standard model gives the same weights to all the assets belonging to a particular class without considering the size of the obligors and their financial stability or their reaction with respect to the market movements.
Anyway, both models does not consider any type of stress so they are not reliable.

4 Capital stress test

4.1 LGD and k modelled as a double t-student

In this last part, we made the same consideration regarding the Regulatory capitals, the Add-on and also the robustness check changing the distributional assumptions of the parameters k and LGD.
In particular, we used the same Montecarlo approach but considering a double t-student with the same parameter μ and σ defined in the first section, considering the degrees of freedom from 2 to 20.

$$k = \overline{k} + \sigma_k(t_1(1,:) \cdot \sqrt{\rho} + t_1(2,:) \cdot \sqrt{1-\rho})$$
$$LGD = \overline{LGD} + \sigma_{LGD}(t_2(1,:) \cdot \sqrt{\rho} + t_2(2,:) \cdot \sqrt{1-\rho})$$

Where both $t_1(2, N)$ and $t_2(2, N)$ are two independent t-students distributions with ν degrees of freedom.
So, we obtained a new random variable, with unknown distribution since the combination of two t-students is not known, with the same mean defined in the previous section and variance equal to $\sigma^2 \frac{\nu}{\nu-2}$ if $\nu > 2$ and not finite mean if $\nu = 2$.
we sampled two different 2 x N matrices following a t-student distribution for both k and LGD, with N = 10^7 simulations, in order to combine them and obtain two different N-dimensional vector for k and LGD.
For both of them we chose $\rho = 0.5$ by simplicity since we wanted to compute a generic double t-student, so we had to choose $0 < \rho < 1$ otherwise we would obtain a t-student, which is a particular case that we analyze in the following section.
Except for the parameters simulations, all the other steps are the same as the Large Homogeneous portfolio case: we stress that we are using again the Vasicek model, so the probability to default, the Expected loss and the quantile computation are the same that we saw in the previous parts, since the idiosinchratic term and the market factor are i.i.d. standard gaussian.

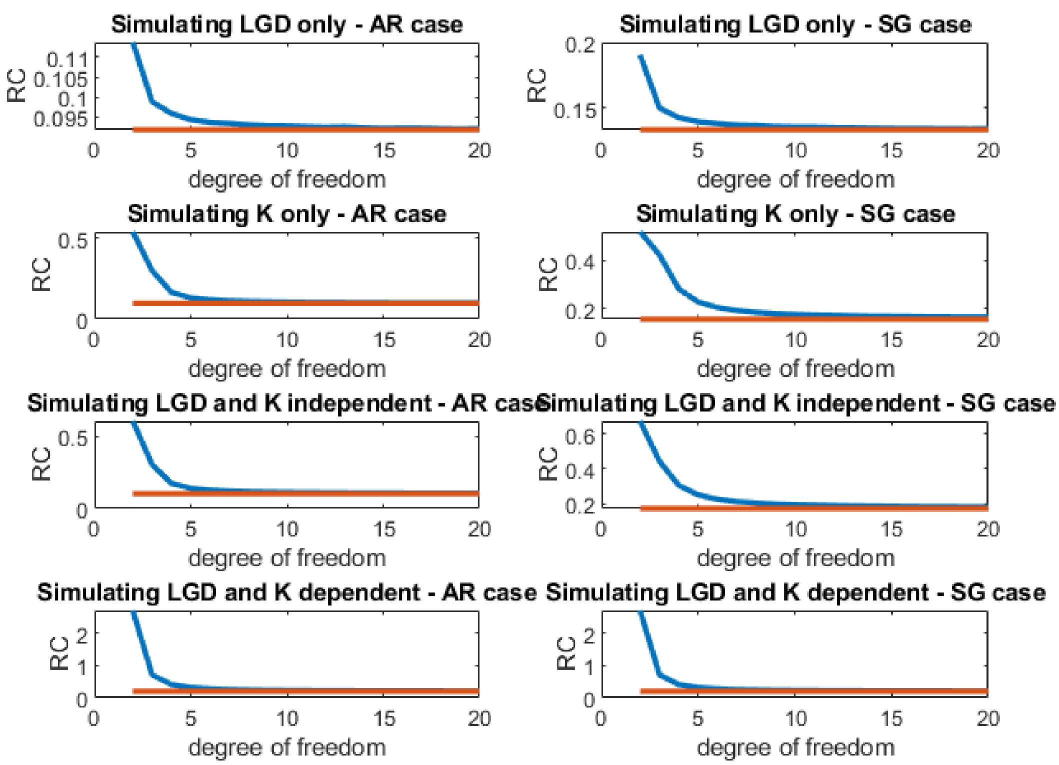


Figure 3: Regulatory capitals for $\alpha = 99.9\%$ in the double t-student case (Red line represents the LHP RC)

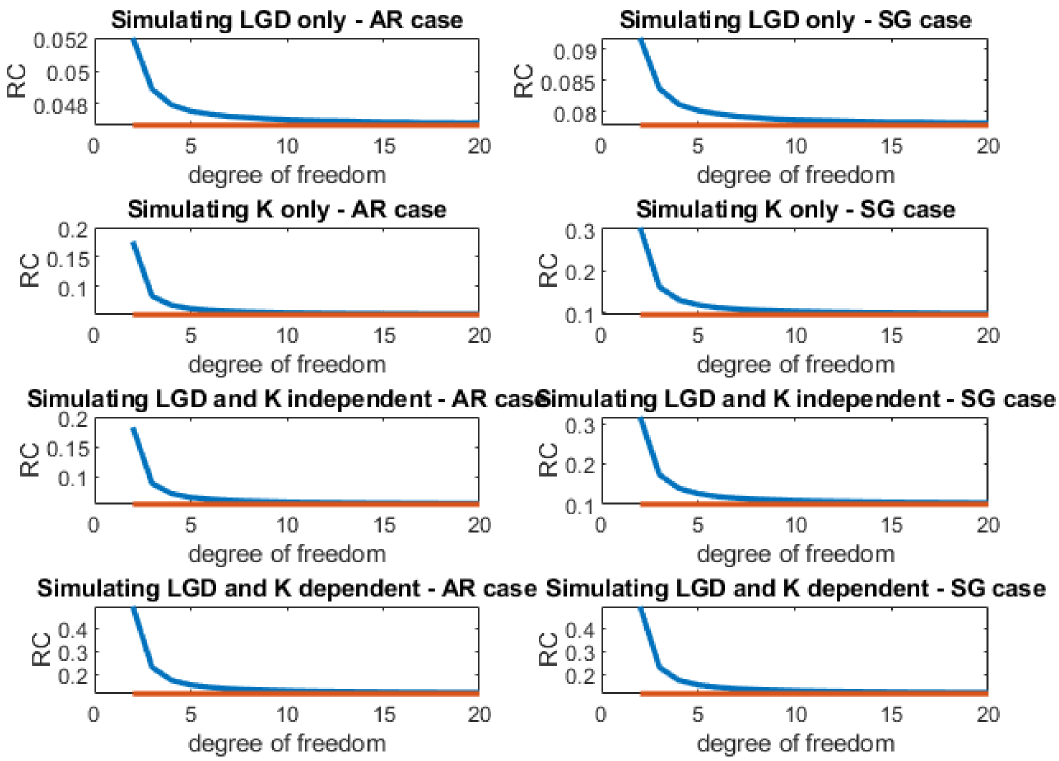


Figure 4: Regulatory capitals for $\alpha = 99\%$ in the double t-student case (Red line represents the LHP RC)

Each simulation for each ν, considering both the values of α and both AR and SG, took approximately a time close to 10 seconds, so the total computation time is almost 180 seconds. We could clearly see that the Regulatory capitals converges to the LHP value, but they never reach it: a double t-student distribution presents fatter tails than a gaussian, so the quantile are always greater than the normal ones.
From the graph we could also notice, from a qualitative point of view, that the RCs are really close after 6 degrees of freedom which seems coherent since the variance of the double t-student is $\frac{3}{2}$ and it start to getting closer to 1.
We expected this result since the t-student converges in law to a gaussian for $\nu \to +\infty$: so a double t-student, which is a combination of two t-students, converges in law to a gaussian too.

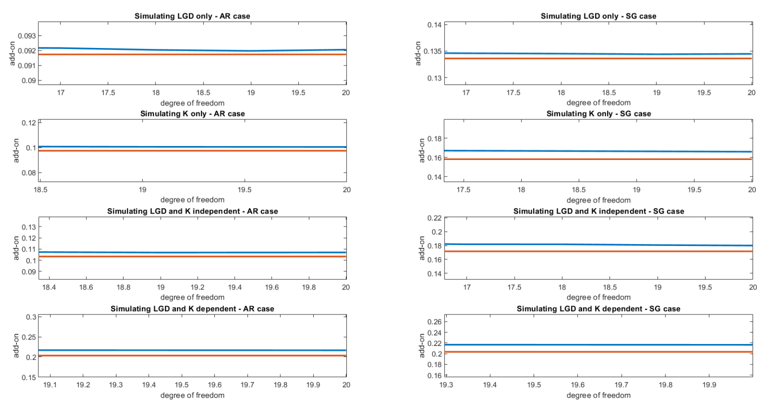


Figure 5: Zoom (dof = 17:20) Regulatory capitals for $\alpha = 99.9\%$ in the double t-student case (Red line represents the LHP RC)

We also computed the HP portfolio case considering $N_{obligors} = 50$ with the same simulations $\mathrm{N} = 3 \cdot 10^5$ and we obtained the following behaviour for $\nu = 2, ..., 20$.

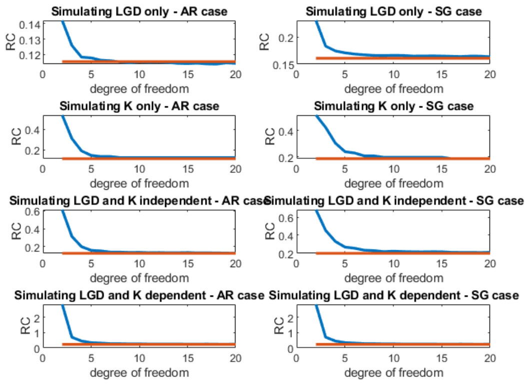


Figure 6: $\alpha = 99.9\%$

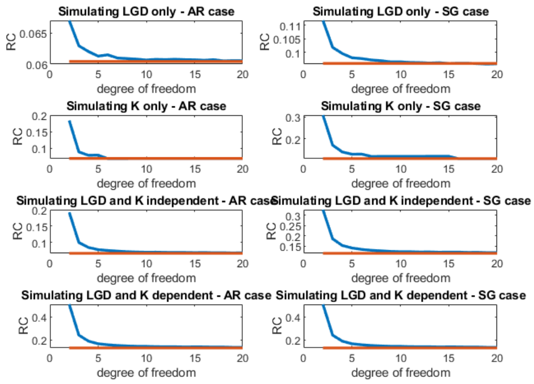
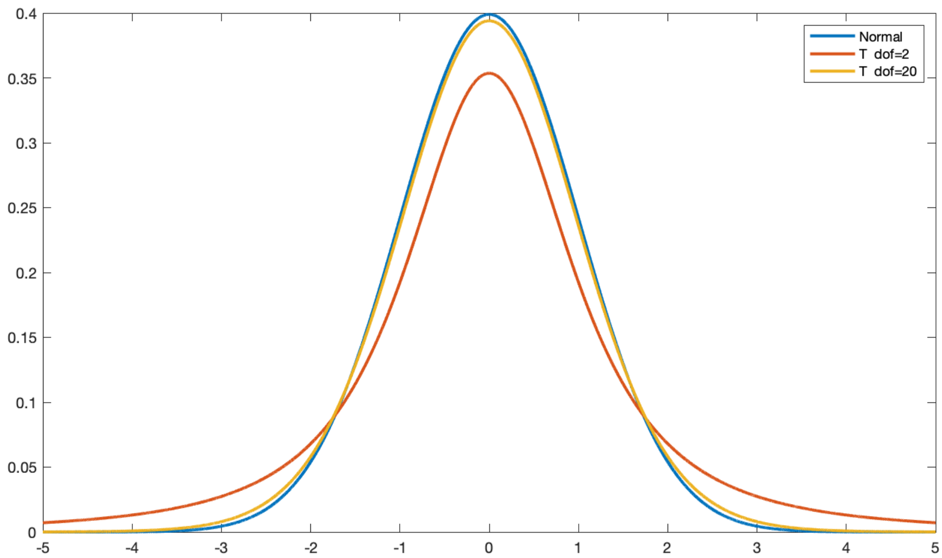


Figure 7: $\alpha = 99\%$

We obtained the same behaviour since the RC converges to the value obtained in the gaussian case with the same finite number of obligors.
Differently from the previous case, for some degrees of freedom the Regulatory capitals are smaller than the asymptotic case value obtained with the gaussian distribution: this is due to the smaller number of simulations. We noticed that the computation time is similar for each simulation, so the total computation time is approximately 10 seconds.
Regarding the Add-ons we obtained the same behaviour of the Regulatory capital obtaining large values for the smallest ν's and very close values for the greatest ones.
We printed them in the code.

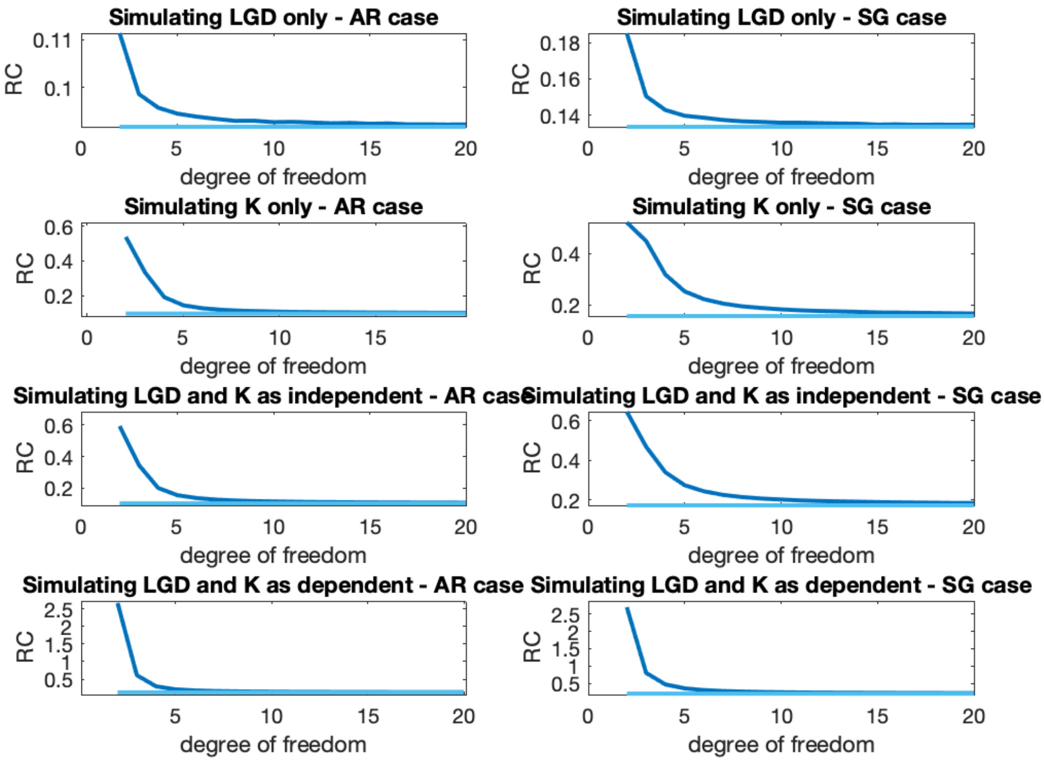
4.2 LGD and k modelled as a t-student

We also conducted a capital stress test by modeling the parameters as t-distributions with degrees of freedom ranging from 2 to 20. Such a simulation means "fattening" the tails of the simulated parameters, especially when the degrees of freedom are low, thereby worsening the cases where the loss is greater (the tails, precisely). We therefore expect a significantly higher Regulatory capital compared to the gaussian case, particularly in simulations where we use low degrees of freedom. The difference between the distributions can be observed in the following picture.



The simulation was performed via Monte Carlo, similarly to the gaussian case, simulating k and LGD distributed as t-Student. The parameters used are those calibrated in Section 2.5.

What we obtained are significantly higher RCs for low degrees of freedom. With 2 dof, we reached an RC more than 10 times greater than the Normal case, which is plausible considering how much the tails fatten in the graph above and the fact that the t-student in this case has infinite variance. As the degrees of freedom increase, we converged, as expected, to the RCs of the Normal case.
The Add-ons follow the same dynamics.
In this section, we simulated using only $\alpha = 99.9\%$, not finding it necessary to add an additional stress factor such as $\alpha = 99.0\%$. Moreover, we did not perform the granularity check since the aim of this part was the comparison between the different distributional assumptions.

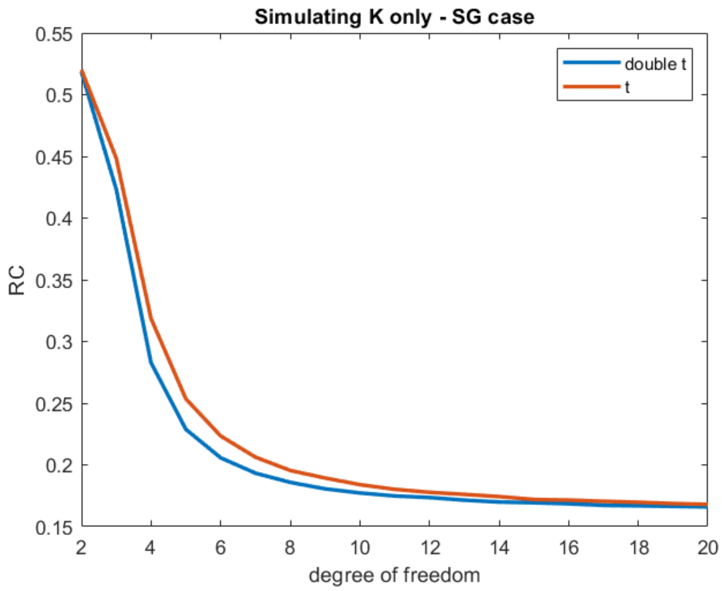


The computational cost is the same of the double t-Student case, for each simulation and for each ν, considering both the values of and both AR and SG, took approximately a time close to 7 seconds, so the total computation time is almost 130 seconds.

4.3 Comments

We just want to make some observation regarding the impact of the fat tails in the Regulatory capitals computation.
For $\nu = 2$ degrees of freedom the computation is not reliable since the variance is not finite and the Regulatory capitals became quite large, while for the other degrees of freedom we noticed the convergence with respect to the LHP case: we already explained why we expected this.
Moreover, the double t-student approximation for the parameters could be consider better than the normal one since the tails of the gaussian are too thin and in case of crisis it does not detect properly the real risks, this holds for both t-student and double t-student case.

We also compared the t-student simulation with the double t-student one: we noticed that they are really close to each other, but the double t-student Regulatory capitals are always lower than the t-student ones except from the smallest values of ν which have a greater volatility and it may happen that this is not verified.
We plot only one particular case which was the most clear (the complete plot is not significant since the curves are really close, anyway it is contained in the Matlab code)



In the t-student case the tails are fatter than both the double t-student and the gaussian case: so, if we want a more cautious approach we can use the t-student or the double t-student distribution for the simulation.
Finally, we could conclude that the last two approach could be better in order to be cautious, but the t-students computation is way more expensive than the gaussian one.

5 Facultative: Python code

A brief description of the Python implemented code is reported here.
We rewrite the code also in Python and it is similar to the Matlab one, but there are some exceptions. First, the Royston test is not available on Python and we used another test, the Henze-Zirkler test, in order to verify the multivariate normality assumption. We obtained way worse results because the p-values are very small.

	p-Value
$k_{SG} - LGD$	0.0038
$k_{AR} - LGD$	0.0052

Since the confidence level was 0.001, we still accepted the null hypothesis since the p-value are higher.

Using the command *fsolve* on Python we obtained slightly different values from the Matlab case for $\overline{k}$.

	$\overline{k}$
SG	-1.7763
AR	-2.2059

We simulated in the same way obtaining slightly different values for every simulation.
We decided to use $N = 3 \cdot 10^6$ and $N = 2 \cdot 10^5$ simulations for, respectively, the LHP cases and the HP cases because the computation time was too high using the ones in the Matlab code.

In general we obtained very similar results and, for each plot, the behaviour is the same described in the report.

Total computation time Python code: 5 minutes and 45 seconds.

References

[1] R. Baviera, 2021. *The measure of model risk in credit capital requirements. Finance Research* Letters, 102064.

[2] J.P. Royston, 1982. *An Extension of Shapiro and Wilk's W Test for Normality to Large Samples. Journal of the Royal Statistical Society. Series C (Applied Statistics) , 1982, Vol. 31, No. 2 (1982), pp. 115-124*

[3] J.P. Royston, 1983. *Some techniques for assessing multivarate normality based on the Shapiro-Wilk W. Journal of the Royal Statistical Society: Series C (Applied Statistics) 32 (2), 121–133.*

[4] D.G. Bonett, T.A. Wright, 2000. *Sample size requirements for estimating Pearson, Kendall and Spearman correlations, PSYCHOMETRIKA–VOL. 65, NO. 1, 23-28.*

[5] G. Loffler, 2003. *The effects of estimation error on measures of portfolio credit risk, Journal of Banking & Finance 27 (8), 1427–1453.*

[6] N. Tarashev, 2010. *Measuring portfolio credit risk correctly: Why parameter uncertainty matters. Journal of Banking Finance 34 (9), 2065–2076.*

[7] G.J. Hahn, W.Q. Meeker, 1991. *Statistical Intervals* (Wiley)