EBA STAFF PAPER SERIES

N.23 – 11/2025

A RATIONALE OF THE PD FLOOR UNDER THE IRB FRAMEWORK

by Simone Casellina, Simone Landini and Mariacristina Uberti

ABSTRACT

The Prudential Regulation has raised the issue of estimation errors due to Internal Rating Based (IRB) estimation process that may produce underestimation of the risk measures. In the context of credit risk, lower bounds (i.e. floors) for the estimated parameters are introduced to limit the impact of such possible underestimation. These floors are heuristically justified by the difficulties to estimate the parameters when the default event becomes rare, as in the case of Low Default Probability Portfolios (LDP). In this paper, on the basis of a standard Asymptotic Single Risk Factor (ASRF) model, and by means of Monte Carlo simulations, we provide a robust justification to PD floors, and a framework for their calibration. Our results give hints that the introduction of a floor can indeed mitigate the possibility that the risk measures become less reliable.

KEYWORDS

Basel 2; Margin of Conservatism; Value-at-Risk; Low default probability; Estimation error; PD floor

JEL CODES

JEL: C13; C15; C54; G17; G21; G28;

1. INTRODUCTION[1]

The Basel Committee on Banking Supervision (BCBS) (2006) introduced a risk-based framework (Basel II) into the system of prudential regulation for banks, allowing them to use internal models to calculate minimum capital requirements for major risk types. Since then, banks may use their estimates of the probability of default (PD), although the Regulator imposed that the "*PD of an exposure shall be at least 0.03%*" (EU, 2013; p. 105)[2]. Such a floor has been introduced in recognition of the difficulty banks would face in estimating and validating PD estimates of this magnitude.

There are various trade-offs regarding the appropriate calibrations of the floors. As stated in BCBS (2016; p. 6) "Floors on individual model parameters can be applied in a targeted way to address concerns about the reliability of particular inputs for particular portfolios. For example, PD floors address the problem that in low-default portfolios, a large number of observations are needed to give confidence in the estimated PD". Moreover, the floors can reduce the variation in model parameters for the same exposure increasing the comparability among banks. Too high floors, however, could bias the risk measures.

Under the revised prudential framework of Basel III, the BCBS (2017) introduced the so-called output floors and reviewed the input floors. These are used to limit excessive variability of the banks' estimates and to increase their reliability. In particular, "*The PD for each exposure that is used as input into the risk weight formula and the calculation of expected loss must not be less than 0.05%.*" (BCBS, 2017; p. 65). This change was implemented because banks did not always have sufficient information on historical default observations. Among the many aspects of the overall framework of Basel III, Regulators believe that input floors are an important aspect to increase the robustness of the IRB approach.

In order to gauge the expected impact of the Basel III reforms, the European Banking Authority (EBA, 2017; 2019) conducted a survey of the major EU banks. In terms of the impact of PD input floors, the study found that the greatest impact in terms of minimum capital requirements will be on the so-called Low Default Portfolios, particularly loans to financial institutions and large corporate customers. Despite the relevance and the expected impact of the PD floor, no clear indication about its quantitative calibration were provided by the BCBS while, as mentioned above, they were heuristically justified by the difficulties to estimate the parameters when the default event becomes rare.

In this paper we study the necessity of the introduction of PD floors. To this end, we specialise the approach suggested by Casellina et al. (2023) in the low default probability case to tackle the problem of the variability of the PD estimate. We provide a robust justification to the floors, and a framework that would enable calibrating them. Setting the minimum level of a parameter that can be reliably estimated is not a so often discussed topic in the literature. Papers such as Pluto and

[1] The opinions expressed are those of the authors and do not involve responsibility of the institutions. We gratefully thank the anonymous reviewers for precious suggestions and comments that helped us to improve the paper.

[2] See Articles 160 (1) and 163 (1) in Section 4 of the of Regulation (EU, 2013) No 575/2013 - known as Capital Requirements Regulation (CRR).

Tasche (2014) and Blümke (2020) deal with providing an estimator of the PD parameter in the presence of a limited number of observed defaults. The topic we deal with in this paper, however, is different. The estimator we consider is the simple ratio between the number of defaults and the number of observations. We ask ourselves up to what minimum value of the PD parameter such an estimator is reliable. To do so, we introduce the problem of the estimation error in the context of the Supervisory model for the quantification of credit VaR, and show that, as the PD parameter decreases, it becomes more difficult to correct the distortion that the estimation error induces on the estimated of VaR.

The rest of the paper is organized as follows. Section 2.1 introduces notation and briefly reviews the framework underlying the IRB approach, i.e., the ASRF model. Section 2.2 addresses the proposed approach to control for the estimation error. Section 3 focuses on the proposed PD floor model to correct the α-quantile estimation so as to set aside the bias induced by the variability of the PD estimator. Section 4 provides Monte Carlo (MC) simulation results to assess the validity of the results obtained for different values of parameters. Section 5 concludes. Appendices contain the meta-code of the developed program listings that the reader may implement with the preferred programming language.

2. The IRB theoretical framework and the PD estimates

2.1 Notation and basic assumptions of the ASFR model

For credit risk, the BCBS relies on a stochastic credit portfolio model aimed at providing the estimate of the loss amount which will not be exceeded with a given confidence level α, that is arbitrarily set, and the corresponding loss threshold is the VaR at this confidence level: VaR_{α}. The VaR estimates the worst-case loss over a target horizon that will not be exceeded with a given level of confidence (Jorion, 2006), i.e., VaR is the α-quantile of the loss distribution. In the credit-risk framework, the confidence level α is commonly set to, at least, 99.5% (Bolder, 2018) but, for the IRB approach, the BCBS sets the confidence level to 99.9%: "*an institution is expected to suffer losses that exceed its level of Tier 1 and Tier 2 capital on average once in a thousand years. This confidence level might seem rather high. However, the high confidence level was also chosen to protect against estimation errors, that might inevitably occur from banks' internal PD, LGD and EAD estimation, as well as other model uncertainties*" (BCBS, 2005; Section 5.1, p.11)[3].

The ASRF model is the baseline for the derivation of the credit risk measures under the IRB approach (Bolder, 2018). Following the classic structural Merton-Vasicek model (Merton, 1974; Vasicek,

[3] It is worth noticing that the IRB approach was originally calibrated for large, internationally active banks and for those banks where an A- rating is typically expected to be needed for a sustainable business model. And A- corresponds roundabout to a historical 0.1% default rate suggesting a 99.9% confidence level.

1987, 1991; Gordy[4], 2003) the creditworthiness change of the i-th exposition is defined as a function of two random variables and a parameter $Y_{i,t} = Y\left(Z_t, W_{i,t};\ \omega\right)$. More explicitly, $Y_{i,t} := \sqrt{\omega} \cdot Z_t + \sqrt{1-\omega} \cdot W_{i,t}$ where $Z_t \sim \mathcal{N}(0,1)$ is assumed a systematic risk factor, that homogeneously spreads its effect on each single borrower, and $W_{i,t} \sim \mathcal{N}(0,1)$ is an idiosyncratic term, that heterogeneously hits the *i*-th borrower only. Parameter $\omega \in (0,1)$ is an exogenous correlation parameter set by the Regulator[5]: portfolios of different instruments have their specific value of ω, also named as the factor loading, that shapes the correlation of the systematic risk factor with the individual creditworthiness change. In the case of large corporate portfolios, the parameter ω ranges between 12% and 24%. It is further assumed that the default event for the *i*-th counterpart is triggered by $Y_{i,t}$ as $D_{i,t} = 1\{Y_{i,t} < s\}$, where s is a given threshold; $1\{\dots\}$ is the standard indicator function that returns 1 if the statement is satisfied or 0 otherwise. The default rate (DR), i.e., the percentage of defaults observed in a given period t, is different from PD because it depends on the value realized by Z_t in that period. By conditioning to a given realization $Z_t = z$ one finds $Y_{i,t}^z \sim N\left(\sqrt{\omega} \cdot z, 1-\omega\right)$. Therefore, the probability of default conditioned to $Z_t = z$ is $\mathbb{E}(DR) = \mathbb{P}\left(Y_{i,t}^z < c\right) = \Phi\left(\frac{c-\sqrt{\omega}\cdot z}{\sqrt{1-\omega}}\right)$, where $c = \Phi^{-1}(PD)$, PD is the long run probability of default[6], $\Phi^{-1}(\cdot)$ denotes the inverse-cumulative Gaussian distribution function, and $\mathbb{E}(DR)$ is the expected value of the default rate distribution. Under this framework, it can be shown that the following expression provides the α-quantile of the default rate distribution[7]:

$$q_\alpha(DR) \equiv VaR_\alpha(PD) := \Phi\left(\frac{\Phi^{-1}(PD)-\sqrt{\omega}\,\Phi^{-1}(1-\alpha)}{\sqrt{1-\omega}}\right) = \Phi\left(\frac{\Phi^{-1}(PD) + \sqrt{\omega}\cdot\Phi^{-1}(\alpha)}{\sqrt{1-\omega}}\right) \tag{1}$$

where the last expression is involved in the BCBS setting to simplify the usage, so when z is evaluated at its worst-case α-level outcome, BCBS replaced $\Phi^{-1}(1-\alpha)$ with $\Phi^{-1}(\alpha)$. This changes the sign in the argument of $\Phi^{-1}(\cdot)$, makes (1) easier to follow, and introduces a negative relationship between the systematic variable and default outcomes implicit in the one-factor Gaussian threshold model.

The quantity $q_\alpha(DR) \equiv VaR_\alpha(PD)$ is the VaR with a confidence level α. In other words, the default rate should exceed this quantity with probability $\mathbb{P}\left(DR > q_\alpha(DR)\right) = 1-\alpha$: that is to say that "exceptions" like $DR > q_\alpha(DR)$ should be materially observed in $(1-\alpha)\%$ of the cases. As mentioned above, under the IRB framework the level of confidence α is set equal to 99.9% so that it is expected that the default rate exceeds the VaR with probability 0.1%.

To compute the VaR it is necessary to estimate some parameters and these estimates are subject to uncertainty due to the estimation process. Replacing the true parameters' value in the theoretical formula (1) with sample estimates, which are based on sampling observations, introduces an additional source of uncertainty, and this implies the so called "estimation risk". In practice, the quantity $q_\alpha(DR)$ is substituted with its plug-in equivalent that can be understood as an estimator (see Appendix C):

[4] Gordy 2003 introduced the portfolio invariance property. This property loosely speaking allows to compute a capital requirement for one exposure without considering the composition of the rest of the portfolio and without assuming that the portfolio is homogenous.

[5] See Article 153 and 154 in Regulation (EU, 2013) No 575/2013 for specific values of ω.

[6] The long run probability of default is the time average of the portfolios' default rate.

[7] As a shorthand notation $q_\alpha(DR)$ represents the α-quantile of the distribution of the quantity in argument.

$$\hat{q}_\alpha(DR) \equiv VaR_\alpha\left(\widehat{PD}\right) := \Phi\left(\frac{\Phi^{-1}(\widehat{PD}) + \sqrt{\omega}\cdot\Phi^{-1}(\alpha)}{\sqrt{1-\omega}}\right) \tag{2}$$

where $\widehat{PD}$ is the estimator of the parameter PD. It is worth noticing that $\hat{q}_\alpha(DR)$ is an unbiased estimator of $q_\alpha(DR)$, i.e., $\mathbb{E}[\hat{q}_\alpha(DR)] = q_\alpha(DR)$.

In Casellina et al. (2023), the impacts of the variability of the PD estimates are analyzed in the IRB theoretical framework. In particular, it is shown that when the parameter PD in (1) is substituted with its estimate $\widehat{PD}$ as in (2), the probability that the default rate exceeds the quantity $\hat{q}_\alpha(DR)$ is higher than $1-\alpha$, say $\mathbb{P}\big(DR > \hat{q}_\alpha(DR)\big) > 1-\alpha$ and this is due to the estimation error that is to be taken into account.

2.2 The proposed approach to control for the PD estimation error

The main issue to be clarified is that the parameter PD is a theoretic notion while, in practice, it is a parameter to be estimated, and this aspect introduces the estimation risk issue. As a consequence, even if the estimator of the input parameter is unbiased, its sample variability can introduce a bias in the VaR measure.

Casellina et al. (2023) propose a computational approach, here described with technical detail in the Appendices, aimed at obtaining a correct estimate of the α-quantile of the default rate distribution correcting for the variability of the estimator $\widehat{PD}$. It is worth stressing that *a point of strength of this approach is that it is completely specified within the regulation framework, without further assumptions*. This approach consists in substituting the estimate of the PD with an appropriate upper bound of a confidence interval estimator (see Appendix D):

$$\widehat{U}(\beta) \equiv u\big(\beta; \widehat{PD}\big) := \widehat{PD} + \Phi^{-1}(\beta)\sqrt{\sigma^2_{\widehat{PD}}} \tag{3}$$

with $\beta \in (0,1)$. This expression is obtained from the same underlying hypotheses of the IRB framework and, in particular, the variance of the estimator $\widehat{PD}$ (i.e., $\sigma^2_{\widehat{PD}}$) need not be estimated as it is derived from the same regulation hypotheses, see Appendix A.[8] Therefore, this approach does not introduce any additional hypothesis and it does not require estimating any additional parameter. The estimator of the adjusted-VaR of the default rate distribution is then (see Appendix D)

$$\hat{q}_{\alpha,\beta}(DR) \equiv VaR^{adj}_\alpha\big(\widehat{PD}, \beta\big) = \Phi\left(\frac{\Phi^{-1}\big(\widehat{U}(\beta)\big) + \sqrt{\omega}\,\Phi^{-1}(\alpha)}{\sqrt{1-\omega}}\right) \tag{4}$$

where value of the β-confidence of the upper-bound interval estimator must fulfil the following condition

$$\mathbb{P}\left(DR > \hat{q}_{\alpha,\beta}(DR)\right) = 1-\alpha \tag{5}$$

[8] More specifically, this means that we do not have to evaluate the standard deviation of historical default rates, which could be correlated, but rather we can estimate the variance as described in Appendix D according to a known result of Bluhm, et al. (2010); see Proposition 2.5.9.

Casellina et al. (2023) show that, in general, the level of the correction is not fixed, namely β *should be higher when the number* T *of observations (years) is smaller, when the asset correlation* ω *is higher or the level of the* PD *is lower*, that is what typically happens with so-called "low-PD" portfolios, e.g., large-corporates portfolios or the best rating grades samples of any other portfolios. The key point of the proposed methodology is that β is the control variable of the problem but it has a natural upper bound in 100%.

3. The need for a PD floor

In LDP credit risk assessment, when the PD is low and the default events are infrequent, it can be expected that the difficulties in estimating the PD increase to such an extent that the bias of the VaR estimator also increases, and its reliability decreases. In this paper we show that there are levels of PD, asset correlations and number of observations (years) used for the estimation of the PD, for which it does not seem possible to correct the estimate of the α-quantile of the distribution of the default rates. In other words, in this paper we claim that *the floor should be set at such a level below which it is no longer possible to correct for estimation error in the quantification of the default rate distribution quantile*.

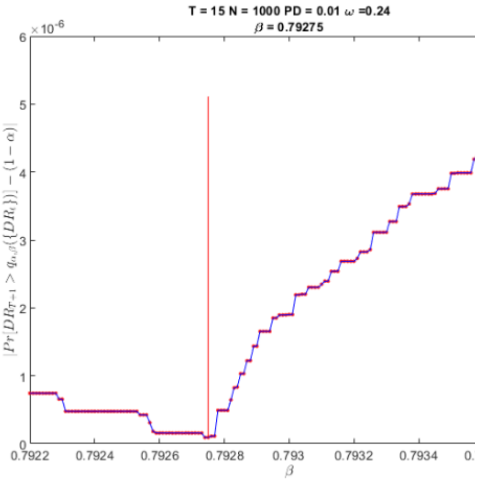
For example, consider the case of a portfolio of $N = 1{,}000$ large exposures, $T = 15$ years of observations, and estimated PD equal to 1%. The Figure 1 represents the outcomes of the numerical process used in Casellina et al. (2023) to set the value of β (on the x-axis), i.e., the recursive grid-search algorithm described in Appendix D. In practice, β is set so that the difference between the effective probability that the default rate exceeds the estimated quantile is equal to $1 - \alpha$. In other terms, β is chosen so to minimize the quantity:

$$S(\beta) \coloneqq \left|\mathbb{P}\left(DR > \hat{q}_{\alpha,\beta}(DR)\right) - (1 - \alpha)\right| \tag{6}$$

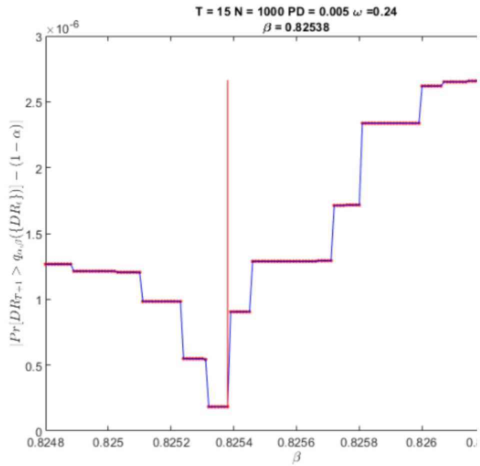
that is represented on the y-axis: notice that $\hat{q}_{\alpha,\beta}(DR) \equiv VaR_{\alpha}^{adj}\left(\widehat{PD}, \beta\right)$ as in (4).

The Figure 1 shows that by setting $\beta = 79.275\%$ it is possible to ensure that the effective probability that the default rate exceeds the estimated α-quantile is equal to $1 - \alpha$ or, in other terms, that $S(\beta) = 0$. Notice that the asset correlation ω is equal to 24% which is the highest level envisaged by the regulation. Moreover, Figure 1 shows that, when the estimated PD decreases, for example from 1% (panel (a)) to 0.5% (panel (b)) and to 0.25% (panel (c)), then it is necessary to increase the level of β to strengthen the level of the correction: from $\beta = 0.79275$ (panel (a)) to $\beta = 0.82538$ (panel (b) to $\beta = 0.86695$ (panel (c)). This outcome suggests that *for extremely low* PD *it will not be possible to correct the estimated quantile of the default rates distribution.*

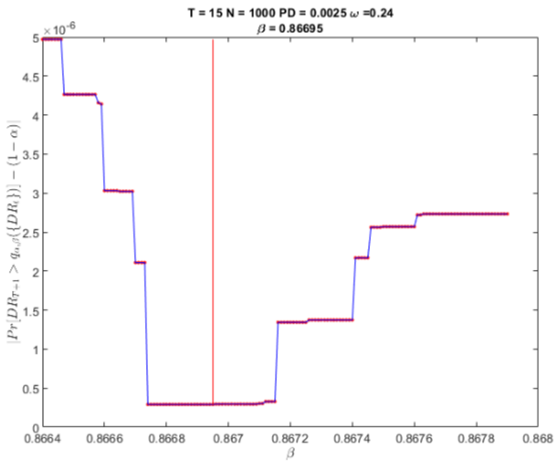
Figure 1: results of the recursive grid-search estimate of β for a portfolio of $N = 1{,}000$, $T = 15$, $\omega = 24\%$ and $PD = 1\%, 0.5\%, 0.25\%$; see Appendix D.



(a) PD=1%



(b) PD=0.5%



(c) PD=0.25%

The optimization is obtained numerically (derivative free algorithm). While situations where the minimum of S(β) corresponds to multiple levels of the PD are possible, in such cases we take prudentially the maximum between the PD values associated with the minimum of S(β). This can be seen in figure 1 (b), but is also clear from the code provided in the Annex: lines 30 - 33 of the Annex D.

4. The PD floor setting

4.1 Introductory example

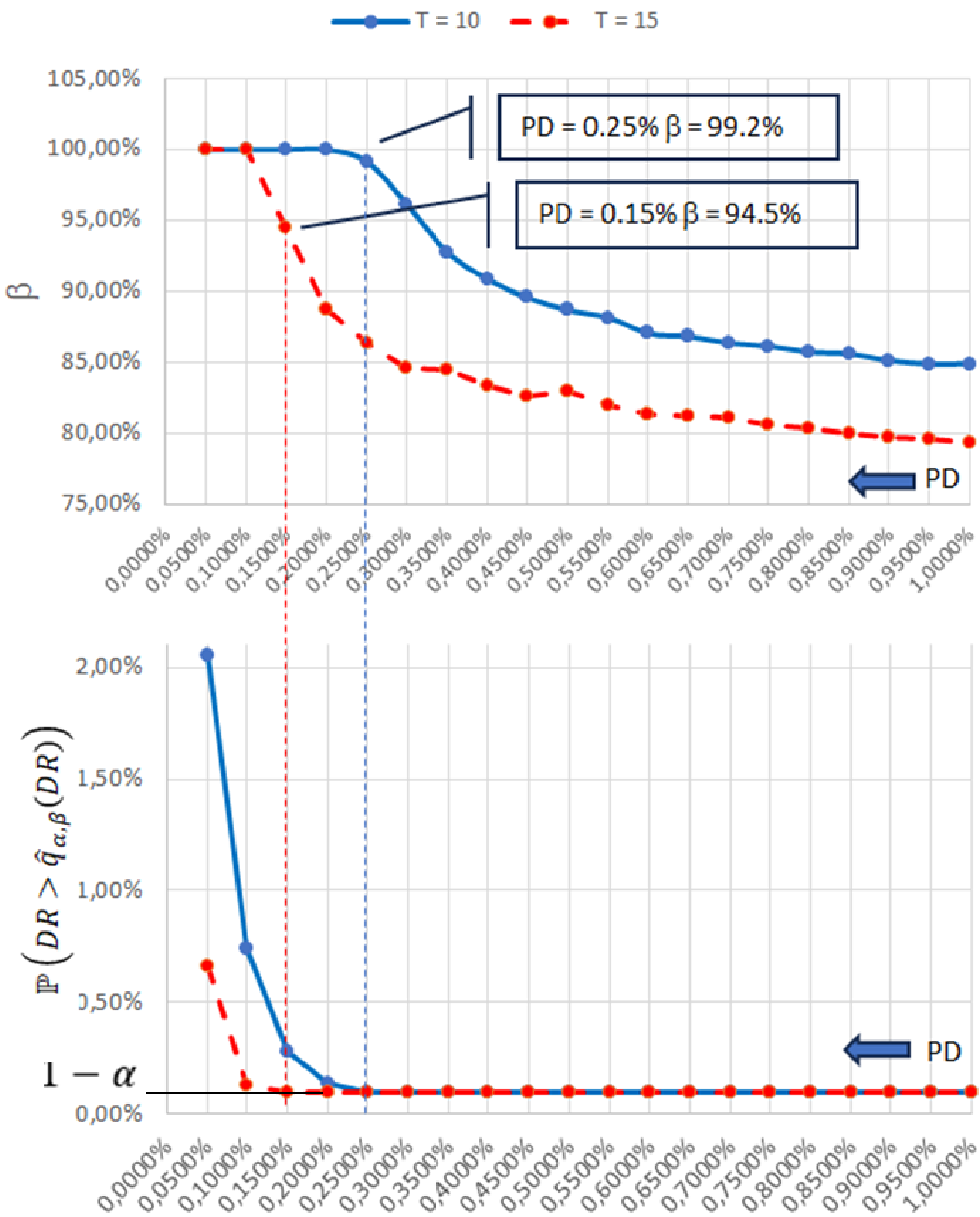
We study the case of portfolios of $N = 1{,}000$ counterparts for which the PD parameter is estimated as the average of T observed portfolio years. Also, we assume we are dealing with large corporates portfolios, then the level of the asset correlation is always $\omega \in [12\%, 24\%]$.[9]

The top panel of Figure 2 showcases how the confidence β increases when the estimated PD decreases from 1% to 0.05% (read the x-axis from right to left). The bottom panel instead shows that the probability that the default rates exceed the estimated α-quantile is exactly equal to $1 - \alpha$, until the estimated PD drops below a certain level. In case $T = 10$ years (solid lines) ten year are used for the estimation, while the lowest PD for which it is possible to correct the estimated α-quantile is equal to 0.25%, with associated level of $\beta = 99.2\%$; below this value the quantity $\mathbb{P}\left(DR > \hat{q}_{\alpha,\beta}(DR)\right)$ becomes higher than $1 - \alpha = 0.1\%$.

For example, with a PD equal to 0.05% (i.e., the floor under Basel III), the quantity $\mathbb{P}\left(DR > \hat{q}_{\alpha,\beta}(DR)\right)$ is equal to 2% i.e., 20 times larger than the desired level $1 - \alpha = 0.1\%$. However, with a larger number of observations, for example $T = 15$ (dashed lines), it is possible to correct the estimated α-quantile for a lower level of the PD. Figure 2 shows that with $T = 15$ it is possible to arrive to a PD level of 0.15% with a confidence $\beta = 94.5\%$. Below such value of PD no correction to the VaR is possible because the confidence level β faster reaches its maximum at 100%.

[9] See Article 153 of Regulation (EU, 2013) No 575/2013.

Figure 2: Estimates of β with $N = 1{,}000$ and $\omega = 24\%$ for different values of PD.



4.2 Floors for different values of parameters

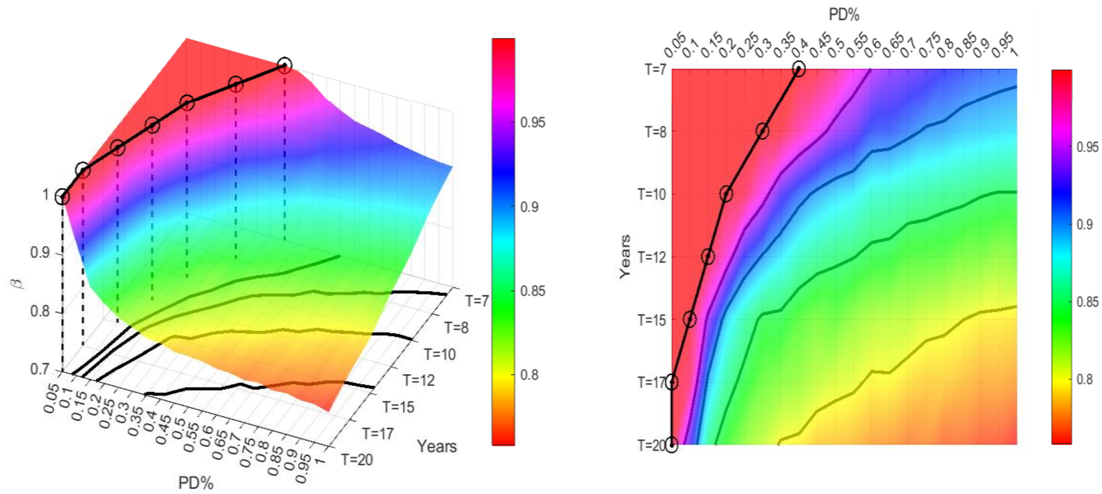
In our approach, the analytic definition of the PD-floor reads as follows:

$$PD_{floor} \coloneqq \arg\min_{PD}\left[S\left(\beta;\widehat{PD}\right) < \ell\,\right]: S\left(\beta;\widehat{PD}\right) = \left|\mathbb{P}\left(DR > \hat{q}_{\alpha,\beta}(DR)\right) - (1-\alpha)\right| \tag{7}$$

Where $\hat{q}_{\alpha,\beta}(DR) \equiv VaR_{\alpha}^{adj}\left(\widehat{PD};\beta\right)$ as in (4) and ℓ is a tolerance level that we have set equal to 0.01%[10]. That is, we search for the lowest PD that fulfils (6) according to the β-confidence adjusted-Var defined in (4).

In Figure 2 we have just seen that for a fixed level of the asset correlation ω and for a fixed size of the portfolio N, there exists a level of the PD, depending on the number, under which it is not possible to correct the measure of the quantile of the default rates distribution. Let us now consider Figure 3 that shows the minimum level of PD (i.e., the floor) for which is possible to correct the estimate of the α-quantile of the default rate distribution as the number of the time observations varies while keeping constant the size of the portfolios and given the same asset correlation $\omega = 24\%$.

Figure 3: Estimates of β with $N = 1{,}000$, $\omega = 24\%$ and different values of T and PD



What

Figure 3 makes clear is that, the shorter the time series of default rates (T) is, the higher the confidence (β) is needed to adjust the VaR, and even higher levels of β are needed to adjust the VaR if the PD gets lower and lower.

[10] In numerical simulations we set $\ell = 0.01\%$. We observed that smaller values do not significantly improve the results, while being more computationally demanding to obtain.

Table 1: PD floors given different portfolio size (N), time series length (T) and correlation (ω)

	T \ N	250	500	750	1000	1500
ω = 24%	7	1.200%	0.700%	0.550%	0.450%	0.350%
	10	0.700%	0.400%	0.300%	0.250%	0.200%
	15	0.400%	0.200%	0.150%	0.150%	0.100%
	20	0.250%	0.150%	0.100%	0.100%	0.075%
ω = 12%	7	0.750%	0.400%	0.300%	0.200%	0.175%
	10	0.450%	0.250%	0.200%	0.150%	0.100%
	15	0.300%	0.150%	0.100%	0.100%	0.050%
	20	0.250%	0.100%	0.100%	0.050%	0.040%

Table 1 provides the minimum level of the PD (i.e. the floor) for which it is possible to correct the estimated α-quantile so that $\mathbb{P}\left(DR > \hat{q}_{\alpha,\beta}(DR)\right) = 1 - \alpha$. This level can be seen as the floor for the PD given the combination of T, N and ω. Notice that the level of the floor obtained is not constant, as it varies with the size N of the portfolio and the number T of years used for the estimation of the PD parameter. In this table, two asset correlation values are considered, which are at the minimum (12%) and at the maximum (24%), as defined by Regulation for large corporates portfolios. The main result is: *the shorter the time series of default rates and the smaller the size of the portfolio, the higher the PD-floor both with the minimum allowed asset correlation and (even more so) with the maximum level of correlation*.

5. Conclusions

In this paper we studied the floor for the PD parameter, introduced by the Regulator in the IRB framework, in the context of credit risk and for low default portfolios: i.e. portfolios characterized by very low probability of default. By means of the Monte Carlo approach we provide a rationale for the PD-floor. The main results of this paper are the following.

First of all, we showed the usefulness of introducing floors on the PD since, below certain values of the long run PD, the impact of the estimation error is such that it is not possible to correct the quantification of the quantile of the default rate distribution.

Secondly, we have shown that the floor level should change based on the number of years used for the estimate, the size of the portfolio and the level of asset correlation. For example, we have seen that a floor of 0.05% (i.e. the regulatory value) may be necessary when 15 years of observations are available and the portfolio (single rating grade in case of calibration by grade) has 1500 positions, i.e. a fairly realistic situation.

Thirdly, we provided a framework that could be used to calibrate the floor on case-by-case basis. For example, we have evidence that with small portfolios (200 obligors), high asset correlation (24%) and a short time-window (7 years) the PD-floor should be definitively higher. The results show that in general, for any combination of portfolio size, asset correlation and time series length there exists a limit for the PD under which it is not possible to adjust for estimation error. Moreover, this limit is not fixed and in particular it decreases with size. As such, it is possible that the PD limit is even lower than the Basel III floor. However, the Basel III floor appears adequate when the portfolio includes at least 1000 borrowers and the time series is longer than 15 years. These appear as normal conditions that can be usually found in practice. As such the Basel III floors appear as justified. These results could also serve to justify the additional requirement of avoiding the construction of excessively granular master scales.

We have not explored the calibration of the floors, as our intention with this paper was to provide a justification for the introduction of the floors. We leave that item for further research, where an analysis of the trade-off between a simple approach like the Basel III one and a more complex system of graduation of floors should also be considered. For example, we cannot exclude that having a single fixed floor or a system of floors would result in practically the same capital requirements.

Our results hint towards the likelihood that the introduction of a floor can indeed mitigate the possibility that the risk measures become less reliable for low default portfolios. We also have highlighted that for small portfolios with low PD observed over a few years, there would be room to set the floor to even higher levels than the one envisaged by the Regulation. This last result, however, depends heavily on the level of the asset correlation ω: lower values of the asset correlation would lead to lower levels of the floors. However, estimating the asset correlation would entail introducing further sources of variability. Extending the analysis to the simultaneous estimation of PD and asset correlation parameters is left for further developments of the here proposed approach.

APPENDICES

The following appendices report the meta-code that can be used for numerical estimation by means of the computational approach of Casellina et al. (2023). It should not be difficult for the reader to adapt this code to the most familiar programming language. Each appendix describes a function, with inputs and outputs, involved in the following algorithm describing the whole procedure.

The algorithm needs of seven parameters to be set: the value of the probability of default $PD = 0.1\%$; the value of the loading parameter $\omega = 24\%$; the value of confidence for the credit VaR $\alpha = 99.9\%$; the size of the portfolio $N = 1{,}000$; the number of years to be simulated $T = 10$; the number of Monte Carlo trials $B = 10{,}000$; $x = 5\%$ gives the fifth percentile of a Gaussian distribution for importance sampling (see Bolder (2018); sec. 8.5). The procedure consists in running the following programs detailed below.

A. Run $\left[\widehat{\boldsymbol{PD}}, \boldsymbol{\sigma}^2_{\widehat{\boldsymbol{PD}}}\right] = est\_PDhat\_and\_Variance(PD, \omega, B, N, T)$
B. Run $[\boldsymbol{DR}_{\boldsymbol{T+1}}, \boldsymbol{W}] = gen\_DR\_Tplus1\_and\_ImpSampWgt(x, \omega, B, N)$
C. Run $\left[\widehat{\boldsymbol{VaR}}_\alpha, X_\alpha\right] = est\_VaRalpha\_and\_Exceptions(\alpha, \omega, \widehat{\boldsymbol{PD}}, \boldsymbol{W})$
D. Run $\left[\hat{\beta}, \hat{\delta}, \widehat{VaR}_{\alpha,\beta}\right] = est\_confBeta\_and\_AdjVaR(\alpha, \omega, \widehat{\boldsymbol{PD}}, \boldsymbol{\sigma}^2_{\widehat{\boldsymbol{PD}}}, \boldsymbol{W}, K)$
E. Compute Monte Carlo estimates but mainly get the estimated PD_{floor}

One may now want to find results for a different parameter setting, for instance enlarging the observation window from $T = 10$ to $T = 15$, or reducing the loading factor from $\omega = 24\%$ to $\omega = 12\%$, and so on. It just takes changing the parameters.

A. Estimation of PD and its variance

With a given parameters setting it simulates B portfolios, as Monte Carlo trials, of N exposures for a period of T years depending on a given value of the PD parameter and of the loading factor ω. It returns $(B \times 1)$ vectors for the estimate $\widehat{PD}$ of the long-run PD (time average of default rates) and its variance $\sigma^2_{\widehat{PD}}$, for each portfolio. Therefore, this function serves to generate data for credit risk analysis, as if a bank observed B portfolios over T years.

function $\left[\widehat{\boldsymbol{PD}}, \boldsymbol{\sigma}^2_{\widehat{\boldsymbol{PD}}}\right] = est\_PDhat\_and\_Variance(PD, \omega, B, N, T)$

```
c = Φ⁻¹(PD) % A PD value assumed as known
for b = 1 : B % b-th portfolio (b-th MC trial)
    % simulates time series for
    for t = 1 : T % t-th year
        z(b,t) ~iid N(0,1) % the systematic factor
        PD^z(b,t) = Φ((c − z(b,t)√ω) / √(1−ω)) % the conditional PD
```

$DR^z(b,t) \sim_{iid} Bin(N, PD^z(b,t))/N$ `% the default rate`

end

$\widehat{PD}(b) = \frac{1}{T}\sum_t DR^z(b,t)$ `% this is the "long-run PD" that estimates parameter` PD

`% if` $\widehat{PD}(b) = 0$ `then` VaR_α `will be` $q_\alpha(b) = 0$

`% hence` $DR_{T+1}(b) > q_\alpha(b)$ `will be an exception`

`% by construction and` $\sigma^2_{\widehat{PD}}(b) = 0$ `by definition`

if $\widehat{PD}(b) > 0$ `% variance: needed to compute the adjustment to the VaR`

`% (eq. (3) and Appendix D)`

$$\sigma^2_{\widehat{PD}}(b) = \left[\Phi_2\left(\Phi^{-1}(DR^z(b)), \Phi^{-1}(DR^z(b)), \begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & \omega\\ \omega & 1\end{pmatrix}\right) - \left(\widehat{PD}(b)\right)^2\right]/T$$

else

$\sigma^2_{\widehat{PD}}(b) = 0$

end

end

B. Generates DR at T+1 and Importance Sampling weights

With a given parameters setting it simulates B portfolios, as Monte Carlo trials, of N exposures with a given loading factor ω at time $T+1$. It returns $(B \times 1)$ vectors of the default rates DR_{T+1} together with importance sampling weights for each portfolio. As long as credit risk analysis and value at risk are needed over a 1-year horizon, this function generates this additional year for each of the previously simulated portfolios. The function also returns Importance Sampling weights for variance reduction to be used in VaR estimation.

function $[\boldsymbol{DR_{T+1}}, \boldsymbol{W}] = gen\_DR\_Tplus1\_and\_ImpSampWgt(x, \omega, B, N)$

$m = \Phi^{-1}(x)$ `% default` $x = 5\%$, `to shift the Gaussian distribution`

for `b = 1 : B % b-th portfolio (b-th MC trial)`

$z_{T+1}(b) \sim_{iid} \mathcal{N}(m, 1)$ `% systematic factor from the shifted Gaussian`

$PD^z_{T+1}(b) = \Phi\left(\frac{c - z_{T+1}(b)\sqrt{\omega}}{\sqrt{1-\omega}}\right)$ `% conditional PD with the shifted systematic factor`

$DR_{T+1}(b) \sim_{iid} Bin(N, PD^z_{T+1}(b))/N$ `% default rate with shifted PD`

$f(b) = \phi(z_{T+1}(b))$ `% Gaussian PDF of the shifted systematic factor`

$g(b) = \phi(z_{T+1}(b) - m)$ `% Gaussian PDF of the rescaled shifted systematic factor`

$w(b) = \frac{f(b)}{g(b)}$ `% Importance Sampling weights`

end

C. VaR_α and exceptions

The inputs are a given confidence α, the estimated PDs with loading factor ω as returned by the Appendix A function $[\boldsymbol{\widehat{PD}}, \boldsymbol{\sigma^2_{\widehat{PD}}}] = est\_PDhat\_and\_Variance(PD, \omega, B, N, T)$, and the importance sampling weights are returned by Appendix B function $[\boldsymbol{DR_{T+1}}, \boldsymbol{W}] = gen\_DR\_Tplus1\_and\_ImpSampWgt(x, \omega, B, N)$. It returns $(B \times 1)$ vectors of B portfolio estimates of the VaR at the chosen α-confidence together with a scalar for the relative share of portfolios for which the probability of the estimated default rate at the $T+1$ exceeds the VaR with the estimated PD.

function $[\boldsymbol{\widehat{VaR}}_\alpha, X_\alpha] = est\_VaRalpha\_and\_Exceptions(\alpha, \omega, \boldsymbol{\widehat{PD}}, \boldsymbol{W})$

```
B = numel(PD̂) % number of simulated portfolio
for b = 1 : B % b-th portfolio (b-th MC trial)
        if PD̂(b) > 0
                q̂_α(b) = Φ((Φ^{-1}(PD̂(b)) + Φ^{-1}(α)√ω) / √(1−ω)) % estimate of VaR_α(PD̂): eq. (2)
        else
                q̂_α(b) = 0
        end
        if DR_{T+1}(b) > q̂_α(b)
                X(b) = w(b) % cases of exception of DR_{T+1}: if PD̂(b) = 0 then
                            % q̂_α(b) = 0, hence DR_{T+1}(b) > q̂_α(b) surely
        else
                X(b) = 0
        end
end
X_α = Σ_b X(b) / Σ_b w(b) % empirical estimate of ℙ{PD < VaR_α} = α or ℙ{PD > VaR_α} = 1 − α.
```

D. Adjusting VaR_α with β confidence: recursive grid-search

Computing VaR_α with confidence $\alpha = \mathbb{P}\{PD < VaR_\alpha\}$ needs knowing the long run PD that can be estimated with $\widehat{PD}$, therefore the estimator $\hat{q}_\alpha$ provides a biased estimate for $VaR_\alpha(\widehat{PD})$, which has an uncertainty due to the estimate $\widehat{PD}$ of parameter PD. To adjust $VaR_\alpha(\widehat{PD})$ we introduce the upper-bound of a confidence interval for $\widehat{PD}$: $\widehat{U}(\beta) = \widehat{PD} + \Phi^{-1}(\beta)\sqrt{\sigma^2_{\widehat{PD}}}$, see (3). This gives $\widehat{VaR}^{adj}_\alpha = \widehat{VaR}_{\alpha,\beta}$ with estimator $\hat{q}_{\alpha,\beta}(DR) = \Phi\left(\frac{\Phi^{-1}(\widehat{U}(\beta)) + \Phi^{-1}(\alpha)\sqrt{\omega}}{\sqrt{1-\omega}}\right)$ of which we need to estimate the confidence β as $\hat{\beta} = \mathbb{P}\{\widehat{PD} < q_{\alpha,\beta}(\widehat{PD})\}$.

The inputs are the B estimates of PD and their variance as returned by Appendix A function function $\left[\widehat{\boldsymbol{PD}}, \boldsymbol{\sigma}^2_{\widehat{\boldsymbol{PD}}}\right] = est\_PDhat\_and\_Variance(PD, \omega, B, N, T)$, the sampling weights as returned by Appendix B function $[\boldsymbol{DR_{T+1}}, \boldsymbol{W}] = gen\_DR\_Tplus1\_and\_ImpSampWgt(x, \omega, B, N)$, the same α-confidence of Appendix C function $\left[\widehat{\boldsymbol{VaR}}_\alpha, X_\alpha\right] = est\_VaRalpha\_and\_Exceptions(\alpha, \omega, \widehat{\boldsymbol{PD}}, \boldsymbol{W})$ and, of course, the same loading factor ω involved in previous functions.

The function returns three scalars: the best (recursively grid-searched) numerical estimate $\hat{\beta}$ of the confidence β needed to adjust $VaR_\alpha(\widehat{PD})$, a tolerance $\hat{\delta}$ of this estimate, and the adjusted-VaR, $VaR_\alpha^{adj}(\widehat{PD}; \beta)$.

function $\left[\hat{\beta}, \hat{\delta}, \widehat{VaR}_{\alpha,\beta}\right] = est\_confBeta\_and\_AdjVaR(\alpha, \omega, \widehat{\boldsymbol{PD}}, \boldsymbol{\sigma}^2_{\widehat{\boldsymbol{PD}}}, \boldsymbol{W}, K)$

```
B = numel(PD̂) % number of simulated portfolios
ε = 10^-9, θ0 = 5 − ε,  θ1 = θ0 + 5 % constants
β0 = 0, k = 1 % initial condition
while(k ≤ K) {% at the k-th iteration it evaluates
    𝔹k = [β(k−1) − θ0 · 10^−k : 10^−(k+1) : β(k−1) + θ1 · 10^−k] % the set of β candidates
    Jk = numel(𝔹k) % number of β candidates in the set
    𝕁k = {j ∈ ℕ | j ≤ Jk} % set of indices identifying the candidates β
    for j = 1 : Jk % for each candidate
        for b = 1 : B % for each simulated portfolio (MC trial)
            if PD̂(b) > 0
                Ûβ(b,j) = PD̂(b) + Φ^−1(β(j)) sqrt(σ²_PD̂(b)) % β(j) = 𝔹k(j) is the j-th
                                                          % candidate at the k-th
                                                          % iteration: eq. (3)
                q̂^DR_α,β(b,j) = Φ( (Φ^−1(Ûβ(b,j)) + Φ^−1(α)√ω) / √(1−ω) ) % VaR_α^adj(PD̂; β) eq. (4)
                if DR_T+1(b) > q̂^DR_α,β(b,j)
                    X(b,j) = w(b) % exception w.r.t. the
                                  % adjusted VaR
                else
                    X(b,j) = 0
                end
            else  % if PD̂(b) = 0 then Ûβ(b,j) = 0 hence
                Ûβ(b,j) = q̂^DR_α,β(b,j) = 0     % q̂^DR_α,β(b,j) = 0 therefore
                X(b,j) = w(b)                   % DR_T+1(b) > q̂^DR_α,β(b,j) is an exception
```

```
            end
        end
        X_{α,β}(j) = Σ_b X(b,j) / Σ_b w(b) % the share of exceptions to the adjusted VaR
        Y_{α,β}(j) = |(1 − α) − X_{α,β}(j)|  % see eq. (6)
    end
```

$\mathbb{J}_k^0 = \left\{n \in \mathbb{J}_k | n = \arg\min_{j \in \mathbb{J}_k}\left[Y_{\alpha,\beta}(j) - \min\{Y_{\alpha,\beta}(j)\}\right]\right\}$ % subset of indices

% identifying the best β at the

% k-th iteration

$u_k = \min \mathbb{J}_k^0$ % index of the smallest best β

$v_k = \max \mathbb{J}_k^0$ % index of the largest best β

$\beta_k = \mathbb{B}_k(u_k)$ % the best estimate of β at the k-th iteration

$\delta_k = \mathbb{B}_k(v_k) - \mathbb{B}_k(u_k)$ % associated tolerance

$k = k + 1$ % updates the iterator

}% while-cycle closes

$\hat{\beta} = \max\{\beta_k\}$ % estimate of β such that $\hat{\beta} = \mathbb{P}\{\widehat{PD} < \widehat{VaR}_\alpha^{adj}\}$

$\hat{\delta} \in \{\delta_k | \beta_k = \hat{\beta}\}$ % tolerance

E. Monte Carlo estimates for a given value of PD with confidence α over B portfolios of N exposures for T years with loading factor ω

Given the outcomes returned by previous functions, Monte Carlo estimates can be computed on the basis of the initial parameters setting

$\widehat{PD}^* = \frac{1}{B}\sum_b \widehat{PD}(b)$ % Monte Carlo estimate of the long run PD

$\hat{q}_\alpha^* = \Phi\left(\frac{\Phi^{-1}(\widehat{PD}^*) + \Phi^{-1}(\alpha)\sqrt{\omega}}{\sqrt{1-\omega}}\right)$ % estimate of $\widehat{VaR}_\alpha$

$\widehat{PD}_\alpha^* = \Phi\left(\frac{\Phi^{-1}(\hat{q}_\alpha^*) + \Phi^{-1}(\alpha)\sqrt{\omega}}{\sqrt{1-\omega}}\right)$ % the Worst-Case-DR given $\widehat{VaR}_\alpha$

$\sigma_{\widehat{PD}^*}^2 = \left[\Phi_2\left(\Phi^{-1}(\widehat{PD}^*), \Phi^{-1}(\widehat{PD}^*), \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \omega \\ \omega & 1 \end{pmatrix}\right) - \left(\widehat{PD}^*\right)^2\right] / T$ % variance of $\widehat{PD}^*$

$\hat{U}_\beta^* = \widehat{PD}^* + \Phi^{-1}(\hat{\beta})\sqrt{\sigma_{\widehat{PD}^*}^2}$ % adjustment for $\widehat{VaR}_\alpha$: upper bound of the confidence interval

$\hat{q}_{\alpha,\beta}^{\widehat{PD}^*} = \Phi\left(\frac{\Phi^{-1}(\hat{U}_\beta^*) + \Phi^{-1}(\alpha)\sqrt{\omega}}{\sqrt{1-\omega}}\right)$ % adjusted VaR, $\widehat{VaR}_\alpha^{adj} = \widehat{VaR}_{\alpha,\beta}$

$PD_{\alpha,\beta}^* = \Phi\left(\frac{\Phi^{-1}\left(\hat{q}_{\alpha,\beta}^{\widehat{PD}^*}\right) + \Phi^{-1}(\alpha)\sqrt{\omega}}{\sqrt{1-\omega}}\right)$ % Worst Case Default Rate given $\widehat{VaR}_\alpha^{adj} = \widehat{VaR}_{\alpha,\beta}$,

% the PD-floor

F. Stability analysis of parameters estimates

In this section, we provide stability analysis results about the estimates of the confidence β. We consider an extreme set up of the main parameters: the confidence $\alpha = 99.9$, a portfolio size of $N = 1{,}000$ expositions, a theoretical $PD = 1\%$, a loading coefficient $\omega = 24\%$ and a time series length of $T = 7$ years. With this setting we dealt with 7 experiments, that differ only with respect to the number of simulated portfolios (Monte Carlo trials) as reported in **Error! Reference source not found.**. Moreover, each experiment has been run 100 times.

Table 2 Stability analysis experiments

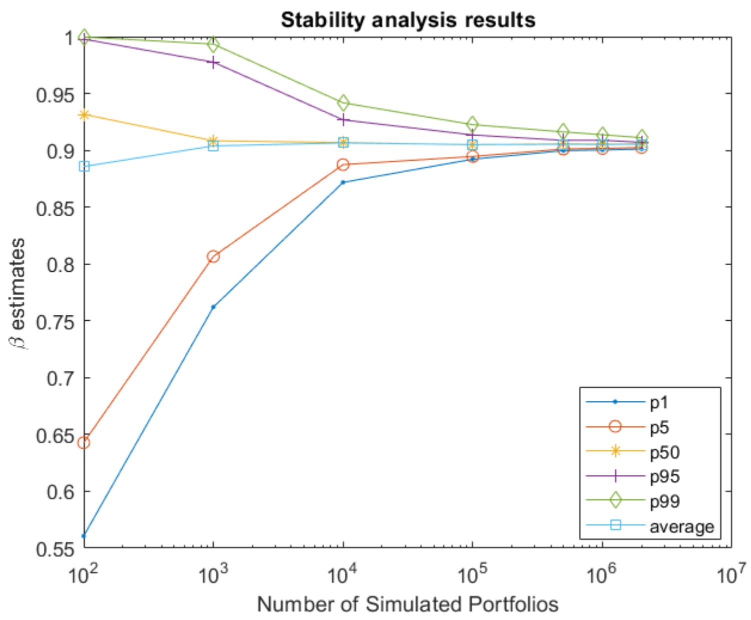
EXP	# MC trials	Min	p1	p5	p50	p95	p99	Max
1	100	0.5431	0.5606	0.6427	0.9316	0.9978	0.9997	0.9997
2	1.000	0.7620	0.7620	0.8066	0.9083	0.9775	0.9934	0.9995
3	10.000	0.8652	0.8718	0.8874	0.9068	0.9267	0.9417	0.9505
4	100.000	0.8920	0.8922	0.8946	0.9049	0.9135	0.9226	0.9290
5	500.000	0.8993	0.8997	0.9011	0.9057	0.9087	0.9162	0.9194
6	1.000.000	0.9001	0.9003	0.9017	0.9054	0.9088	0.9137	0.9168
7	2.000.000	0.9002	0.9010	0.9024	0.9057	0.9070	0.9111	0.9142

EXP	# MC trials	range95	range99	average	stderr	outliers	range99/ average	range99/ p50
1	100	0.3551	0.4390	0.8857	0.0116	0.0000	49.57%	47.13%
2	1,000	0.1709	0.2314	0.9039	0.0050	0.0000	25.60%	25.47%
3	10,000	0.0394	0.0700	0.9065	0.0013	0.0000	7.72%	7.72%
4	100,000	0.0189	0.0304	0.9050	0.0006	0.0000	3.36%	3.36%
5	500,000	0.0076	0.0165	0.9052	0.0003	0.0000	1.82%	1.82%
6	1,000,000	0.0072	0.0134	0.9051	0.0002	0.0000	1.48%	1.48%
7	2,000,000	0.0046	0.0101	0.9053	0.0002	0.0000	1.11%	1.11%

The first part of **Error! Reference source not found.** reports the main sample quantiles of the empirical distribution of the β estimates in each experiment. As the number of trials increases, the variability of the results is reduced, and quantiles stabilise with a narrow difference between the minimum and the maximum: from $[0.5431{,}0.9997]$ in Experiment 1 with 100 replicates to $[0.9002{,}0.9142]$ in Experiment 7 with 2,000,000 replicates.

To better visualise this outcome, consider **Error! Reference source not found.** and the second part of **Error! Reference source not found.** where the $range99$, which evaluates the difference between the 99-th and the 1-st percentiles, is reported. With 100 portfolios, the $range99$ is about 0.4390; with $500{,}000$ portfolios it decreases to 0.0165; finally, it reaches 0.0101 if $2{,}000{,}000$ portfolios are simulated. This means that our procedure provides stable estimates of the confidence β that converges to an average value $\beta = 0.9053$ with a small standard error of about 0.02% and a less than 0.001% of outliers in each experiment, i.e., an exception can be observed once in a million. Finally, the ratios of the $range99$ to the average or to the median ($p50$) become equivalent even with a small number of simulated portfolios (e.g., $1{,}000$), and they are significantly reduced as far as the number of simulated portfolios increases. With $2{,}000{,}000$ simulated portfolios, the $range99$ is about 1.11% of the average or the median. In other words, the empirical estimates of the confidence β converge what can be considered the "true" value with the specified parameters' setting.

Figure 4 Stability analysis results



6. REFERENCES

Basel Committee on Banking Supervision (BCBS). (2005, July 05). *An Explanatory Note on the Basel II IRB Risk Weight Functions.* Tratto da https://www.bis.org/bcbs/irbriskweight.htm

Basel Committee on Banking Supervision (BCBS). (2006, June 30). *Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version.* Tratto da https://www.bis.org/publ/bcbs128.htm

Basel Committee on Banking Supervision (BCBS). (2016, March). *Reducing variation in credit risk-weighted assets - constraints on the use of internal model approaches.* Tratto da https://www.bis.org/bcbs/publ/d362.htm

Basel Committee on Banking Supervision (BCBS). (2017, December). *Basel III: Finalising post-crisis reforms.* Tratto da https://www.bis.org/bcbs/publ/d424.htm

Bluhm, C., Overbeck, L., & Wagner, C. (2010). *Introduction to Credit Risk Modeling* (2nd ed. ed.). New York: Chapman and Hall/CRC. doi:10.1201/9781584889939.

Blümke (2020) Estimating the probability of default for no-default and low-default portfolios, Journal of the Royal Statistical Society: Applied Statistics 69, Part 1, pp. 89–107.

Bolder, D. J. (2018). *Credit-Risk Modelling: Theoretical Foundations, Diagnostic Tools, Practical Examples, and Numerical Recipes in Python.* Cham: Springer International Publishing.

Casellina, S., Landini, S., & Uberti, M. (2023). Credit risk measures and the estimation error in the ASRF model under the Basel II IRB approach. *Communications in Nonlinear Science and Numerical Simulation, ISSN 1007-5704*(118), 17. doi:https://doi.org/10.1016/j.cnsns.2022.106977

European Banking Authority (EBA). (2017). *Guidelines on PD estimation, LGD estimation and the treatment of defaulted exposures.* Tratto da https://www.eba.europa.eu

European Banking Authority (EBA). (2019). *Basel III reforms: impact study and key recommendations.* Tratto da https://www.eba.europa.eu/risk-analysis-and-data/quantitative-impact-study/basel-iii-monitoring-exercise/call-for-advice

European Union (EU). (2013, June 26). *Prudential requirements for credit institutions and investment firms and amending Regulation (EU), Official Journal of the European Union, Regulation (EU) No 575/2013, European Parliament and Council.* Tratto da http://data.europa.eu/eli/reg/2013/575/oj/eng

Gordy, M.B., (2003)m, A risk-factor model foundation for ratings-based bank capital rules, *Journal of Financial Intermediation, 12(3), 199–232*

Jorion, P. (2006). *Value at Risk: The New Benchmark for Managing Financial Risk* (3rd Ed. ed.). New York: McGraw Hill Professional.

Merton, R. C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *The Journal of Finance, 29*(2), 449–70. doi:10.2307/2978814

Katja Pluto & Dirk Tasche, 2006. "Estimating Probabilities of Default for Low Default Portfolios," Springer Books, in: Bernd Engelmann & Robert Rauhmeier (ed.), The Basel II Risk Parameters, chapter 0, pages 79-103, Springer.

Vasicek, O. (1987). Probability of Loss on Loan Portfolio. SAN FRANCISCO, CALIFORNIA, USA.: KMV Corporation. Vasicek's results were published in ‘‘Loan Portfolio Value’’, Risk, 15(12), December 2002, 160-2. Tratto da https://www.moodysanalytics.com/-/media/whitepaper/before-2011/02-12-87-probability-of-loss-on-loan-portfolio.pdf

Vasicek, O. A. (1991). *Limiting Loan Loss Probability Distribution.* SAN FRANCISCO, CALIFORNIA, USA,: KMV CORPORATION.

ACKNOWLEDGEMENTS

We are grateful to participants in EURO 2022 Conference, organized by Aalto University to be held in Espoo, Finland, on July 3–6 2022 for fruitful discussions. This work was supported by the Ministero dell'Università e della Ricerca (MUR), Italy. The authors declare they have no conflict of interest, and their opinions do not involve responsibility of their institutions.

Simone Casellina:
European Banking Authority, 20 avenue André Prothin CS 30154, 92927 Paris La Défense CEDEX, France; https://eba.europa.eu/; SimoneEnrico.Casellina@eba.europa.eu (max. 2 lines)

Simone Landini:
IRES Piemonte, Via Nizza 18, 10125, Torino, Italy; landini@ires.piemonte.it

Mariacristina Uberti:
Full Professor; Department of Management, University of Torino, Corso Unione Sovietica, 218 Bis, 10134 Torino, Italy; mariacristina.uberti@unito.it

eba European Banking Authority

Tour Europlaza, 20 avenue André Prothin CS 30154
92927 Paris La Défense CEDEX, FRANCE
Tel. +33 1 86 52 70 00

E-mail: info@eba.europa.eu

https://eba.europa.eu

ISBN 978-92-9407-208-5
ISSN 2599-7831

doi:10.2853/9429166
DZ-01-25-096-EN-N