Statistical Uncertainty of PD Estimation under the Basel Regulations

Dominik Scherer [1] and Pascal T. Kleinen [2] and Gerrit Reher [2]

[1] Deloitte GmbH Wirtschaftsprüfungsgesellschaft, Germany, doscherer@deloitte.de
[2] Deloitte GmbH Wirtschaftsprüfungsgesellschaft, Germany

January 22, 2023

Abstract

Banks calculating capital requirements for credit risk based on internal ratings according to the Basel framework must add a margin of conservatism to their estimates of probability of default (PD). This margin shall, at a minimum, cover the statistical uncertainty related to the estimation. A wide range of methods and assumptions are regularly used to quantify this uncertainty and these methods and assumptions are frequently challenged and criticized by supervisory authorities. In this article we show why one should distinguish between two different types of statistical uncertainty in PD estimation and between two different approaches to quantify them. We derive formulas for both types and approaches from the credit-portfolio-model assumptions which underlie the Basel risk weight formulas. By a numerical evaluation and simulation we find that, depending on the portfolio and availability of historical data, both types can yield sizeable contributions to the overall uncertainty of a PD estimate. Consequently we discuss the impact of these findings on capital requirements.

Keywords: Probability of Default (PD); Statistical Uncertainty of PD; Estimation Error; Margin of Conservatism; IRB Approach; MoC C.

Disclaimer: Any views expressed in this article are solely those of the authors and not necessarily those of Deloitte GmbH Wirtschaftsprüfungsgsellschaft. Any decision on the basis of information given herein is made exclusively on own risk.

1 Introduction

In the internal ratings-based (IRB) approach to calculate capital requirements for credit risk, banks calibrate their probability of default (PD) models in a through-the-cycle (TTC) manner. The target level of calibration for a rating grade or calibration segment is its long-run average of default rates. For banks in the European Union, the European Banking Authority (EBA) has issued Guidelines [EBA, 2017] which, inter alia, prescribe a way to calculate this long-run average from historical default data and also demand for the estimation of a Margin of Conservatism (MoC) reflecting uncertainties in this calculation.

While banks use various approaches for this estimation of MoC, these approaches are currently under intense scrutiny in internal-model inspections by European supervisory authorities and are a frequent source of criticism in these inspections. Therefore it is important for banks to have a clear understanding of the sources of uncertainties in their PD quantification and of their respective importance. In art. 36 and 42 of [EBA, 2017], EBA defines three categories of MoC:

- MoC related to data and methodological deficiencies (category A);
- MoC related to relevant changes in underwriting standards, risk appetite, collection and recovery policies and to any other source of additional uncertainty (category B);
- MoC related to the general estimation error (category C).

In this article we study the rather generic category C, covering the unavoidable statistical uncertainty present in any PD estimation, where the individual data and representativeness issues of a particular bank do not play a role.

Past experience of the authors in IRB modelling projects at various banks shows that in practice often a one-dimensional approach is chosen to quantify statistical uncertainty of PD, i.e. methods assessing only one source of uncertainty. This source is then, although it is often not stated explicitly, either the finite number of borrowers in the samples from which PD is estimated or the finite length of the data history used for the estimation. In the first case correlations between borrower defaults are often ignored while in the second case they are essential, driving the variations of the observed default rates over time.

In this article we show that both are relevant sources of uncertainty and discuss how to distinguish between them. We address the question of how both can be quantified consistently, exploiting the model assumptions which underlie the Basel risk-weight formulas, taking into account a macroeconomically-driven correlation between borrower defaults. By a numerical evaluation and simulation we find that, depending on the number of borrowers and the importance of the correlation in the portfolio under consideration, both types can yield sizeable contributions to the overall uncertainty of PD and consequently on banks' capital requirements. However we conclude that, while from a purely mathematical point of view the case is clear, further guidance is needed concerning the regulatory relevance of the second type of uncertainty for MoC C under the IRB regulations.

2 Materials and Methods

2.1 Regulatory statements on the PD estimation method

Under European IRB regulations there are no predefined methods to calculate the general estimation error of a risk parameter. Art. 43 (b) of [EBA, 2017] states that the MoC of category C should "*reflect the dispersion of the statistical estimator*". This estimator is, in the case of a through-the-cycle PD estimation for the IRB approach, the average of observed annual default rates over a multi-year period containing economically "good" and "bad" years.

In this article we use the term *long-run average default rate* of a rating grade or portfolio segment for this random quantity, which under the IRB approach is defined over a *fixed* period in time $t = 1, .., T$, and denote it with R_L:

$$R_L = \frac{1}{T} \sum_{t=1}^{T} R_D(t). \tag{1}$$

$R_D(t)$ are the (obviously random) annual default rates of the grade or segment in the time window chosen for the estimation. According to art. 82 of [EBA, 2017] this time window should end at the date of the most recent available portfolio snapshot for which an annual default rate can be calculated and its length should be chosen in such a way that it reflects the "*likely range of variability of default rates*". The observed value of R_L, calculated from the observed default rates $r_D(t)$ is denoted as

$$r_L = \frac{1}{T} \sum_{t=1}^{T} r_D(t). \tag{2}$$

In practice we usually have one single observation r_L per grade or segment and cannot repeat this random experiment. This observation is then the estimate for the PD of the grade or segment.

It is crucial to keep in mind that, from a mathematical point of view, we are estimating a hidden (unobservable) borrower property "TTC PD", using an average of *past* default rates as an estimator. The resulting estimate is then used to calibrate the prediction model providing the risk differentiation between borrowers to the desired level of risk. Since this model is applied to predict *future* default risk this procedure implies that the borrower property "TTC PD" is implicitly assumed to be constant in time.

2.2 Definition and sources of statistical uncertainty in PD

In line with the point of view formulated by the European Central Bank (ECB) in paragraph 140 of [ECB, 2019], we define the statistical uncertainty of the risk parameter

PD as the uncertainty which can be quantified via the distribution of its estimator, and which is driven mainly by the size of the sample on which the estimation is performed (i.e. the number of borrowers and the number of defaults per snapshot) and the length of the timeframe (i.e. the number of snapshots) during which the sample is observed.

Our approach to quantify this statistical uncertainty is to calculate a confidence interval via the distribution of the estimator. The shape of the distribution must be derived from some theoretical assumptions. For the sake of transparency and focus we restrict ourselves, for any analytical calculations, to the simple and well-known Wald confidence interval which implies approximating the binomial distribution of default events by a normal distribution, although this interval is known to have shortcomings in case of a low number of borrowers and default probabilities close to 0 or 1 [Brown et al., 2001]. In the simplest case, which is under the theoretical assumption of uncorrelated defaults, this interval at confidence level $1-\alpha$ can be calculated with a classical textbook formula for a grade or segment of N borrowers:

$$R_{D,\alpha}^{u,l} = r_D \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\frac{r_D \cdot (1 - r_D)}{N}}. \tag{3}$$

for the upper bound $R_{D,\alpha}^{u}$ and the lower bound $R_{D,\alpha}^{l}$ of the interval at any point in time t. Note that there is no time dependence in this formula because given the assumption of uncorrelated defaults one usually also assumes that there is no significant variation of default rates over time if N is sufficiently large.

These assumptions are, however, in conflict with historical experience and also with the basic assumption behind the IRB risk-weight formulas, i.e. that all borrowers are subject to a systematic risk factor, leading to macro-economically driven variations of their default rates over time. This means that the estimator R_L is a linear combination of estimators $R_D(t)$ for the PDs of borrowers under different macroeconomic conditions. These estimators carry different statistical uncertainties. Hence, one question to be answered in this article is:

1. How can these different uncertainties be combined in a way consistent with the assumptions of the IRB approach?

The second question in the focus of our work is:

2. Does the selection of the reference dates t contribute to the statistical uncertainty of PD estimation and, if yes, how can we quantify this type of uncertainty in a way consistent with the IRB assumptions?

In the following, we will address these questions in the framework of the Asymptotic Single Risk Factor (ASRF) model [BCBS, 2005], also known as Vasicek model [Vašiček, 2002], which the regulators have chosen as the mathematical foundation of IRB capital requirements. However, despite the model being "asymptotic" in the sense of maximizing portfolio granularity, we will allow the number of borrowers in a portfolio to be a finite number N to be able to discuss the following two kinds of statistical uncertainty that exist when PD is estimated from historical data:

1. Uncertainty due to the **limited sample size** $N(t)$ in historical portfolio snapshots: This uncertainty is expected to decrease with increasing $N(t)$ and increasing number of snapshots T;

2. Uncertainty from the **choice of the timeframe**: It is uncertain how close the average annual default rate in a randomly chosen timeframe $t = 1, .., T$ is to the "true" TTC PD. This uncertainty is expected to decrease with increasing T, but is not expected to vanish even for an infinitely large portfolio.

These two types of uncertainties have also been identified in [Garcia-Cespedes and Moreno, 2016], however with a focus on the uncertainty in the tail of the portfolio loss distribution, while we focus on the uncertainty of the PD itself.

The two types of uncertainties are directly related to the two questions defined above. As we will see in section 3.1, the first type of uncertainty can be quantified independently for each t in the ASRF model via the conditional variance of R_D, i.e. given a certain macroeconomic situation, and then aggregated to calculate the uncertainty of the long-run average default rate.

Considering the second type of uncertainty, there are two possible ways to deal with it:

- Consider the choice of the timeframe $t = 1, .., T$ as random sampling. In this case, it contributes to the overall uncertainty and can be quantified in the ASRF model together with the first type of uncertainty through a calculation of total variance, see section 3.2.

- Consider the timeframe $t = 1, .., T$ as given, as there are regulatory prescriptions for this choice and a bank has to prove that this timeframe covers a sufficient amount of good and bad years. In this case, there is only uncertainty *conditional* on the known states of the economy at t such that statistical uncertainty is fully covered by the first type mentioned above.

While from a mathematical point of view both types clearly exist, the relevance of the second type for the regulatory Margin of Conservatism in the IRB approach is in our opinion not clarified yet and will be a debate between banks and their regulators.

To conclude the section on materials and methods, the following subsection gives a brief introduction to the ASRF model assumptions before we exploit this model to calculate statistical uncertainties of PDs. We will derive analytical formulas for the two types of uncertainties in sections 3.1 and 3.2. In section 4 we study the numerical relevance of the uncertainties in various situations.

2.3 Framework: The ASRF model

We model borrower defaults by Bernoulli variables $d_i(t)$ with $i = 1, 2, \dots, N(t)$. For simplicity we will assume non-overlapping 12-month periods after each t for the observation of defaults. In the following, we briefly introduce the ASRF model framework.

Assume a granular and homogeneous portfolio where the idiosyncratic probability of default of any borrower is p. Each borrower's asset value is assumed to follow a geometric Brownian motion. At time t it can be written as a function of the standard normal random variable

$$a_i(t) = \sqrt{\rho}\epsilon(t) + \sqrt{1-\rho}x_i(t) \tag{4}$$

with the asset correlation parameter ρ. The random variables

$$x_i(t) \sim \mathcal{N}(0,1) \ , \ \epsilon(t) \sim \mathcal{N}(0,1) \tag{5}$$

are standard-normally distributed and uncorrelated with one another. Any correlation of borrowers beyond the dependence on a common systematic factor is neglected in this model. For a discussion of the consequences of a more explicit type of correlation, see [Miao and Gastwirth, 2004]. Moreover, the time series of $x_i(t)$ and $\epsilon(t)$ are assumed to contain no autocorrelation. The systematic factor $\epsilon(t)$ influences all borrowers' asset values while $x_i(t)$ is an idiosyncratic risk factor for borrower i. Intuitively, a high value of $\epsilon(t)$ corresponds to a "good" state of the economy while a low (negative) value is the opposite.[1]

A borrower is assumed to default when his asset value falls below his contractual obligations. This can be expressed by a default threshold for the variable $a_i(t)$:

$$d_i(t) = \begin{cases} 1 & \text{if } a_i(t) \leq \Phi^{-1}(p) \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

where Φ^{-1} is the inverse cumulative standard normal distribution function. For ρ=0, the expectation value of $d_i(t)$, to be interpreted as borrower PD, is per definition

$$\mathrm{E}(d_i(t)) = P(x_i(t) \leq \Phi^{-1}(p)) = \Phi(\Phi^{-1}(p)) = p \,, \tag{7}$$

independently of the evolution of $\epsilon(t)$. Therefore the expection value of a default rate $R_D(t)$ is also

$$\mathrm{E}(R_D(t)) = p \text{ , given } \rho = 0. \tag{8}$$

For the general case with correlations $\rho \neq 0$, we can calculate a *conditional* expectation value (in the following denoted with a subscript c) for a default rate, i.e. conditional

[1] From a macroeconomic perspective one would naturally assume autocorrelation at least for the time series $\epsilon(t)$. However, the use of the ASRF model lies mainly in studies of the overall through-the-cycle distribution of default rates where the ordering of good and bad years is of minor importance.

on the realization of $\epsilon(t)$ at a given time t:

$$\begin{aligned}\mathrm{E}_c(R_D(t)) &= \mathrm{E}(R_D(t)|\epsilon(t)=\epsilon_t) = \mathrm{E}_c\left(\frac{1}{N(t)}\cdot\sum_{i=1}^{N(t)} d_i(t)\right) \\ &= \mathrm{E}_c(d_i(t)) = P\left(x_i(t) \leq \frac{\Phi^{-1}(p)-\sqrt{\rho}\epsilon_t}{\sqrt{1-\rho}}\right) \\ &= \Phi\left(\frac{\Phi^{-1}(p)-\sqrt{\rho}\epsilon_t}{\sqrt{1-\rho}}\right).\end{aligned} \tag{9}$$

Obviously it is equal to the conditional expectation value for the default variable of every single borrower. It is possible to calculate a long-term expectation value of the default rate by averaging over the conditional expectation values. Remembering that $\epsilon(t)$ is standard-normally distributed and using a Gaussian integral table we find:

$$\mathrm{E}(R_D) = \mathrm{E}(\mathrm{E}_c(R_D(t))) = \int \Phi\left(\frac{\Phi^{-1}(p)-\sqrt{\rho}\epsilon_t}{\sqrt{1-\rho}}\right)\mathcal{N}(\epsilon_t)d\epsilon_t = \Phi(\Phi^{-1}(p)) = p \tag{10}$$

3 Results

As discussed in section 2.2 there are two approaches to calculate the statistical uncertainty of a PD estimation based on historical portfolio snapshots. Either the snapshot dates t are considered being randomly chosen or they are considered as a pre-defined or deliberately chosen set. In the following subsections we derive analytical formulas for confidence intervals under these two approaches in the ASRF model.

3.1 Uncertainty in case of a pre-defined set of portfolio snapshots

In this subsection we consider portfolio snapshot dates t in the past with given realizations $\epsilon(t) = \epsilon_t$ of the systematic risk factor in the ASRF model. Hence, we are calculating the statistical uncertainty of our PD estimation in terms of a confidence interval derived from the *conditional* variance of default rates, i.e. conditional on the states of the economy defined by ϵ_t which are considered relevant for the estimation of the TTC PD. In this case, statistical uncertainty only stems from the fact that we have a finite sample size N and the inherent randomness of individual borrower defaults.

3.1.1 Conditional default rate for a single portfolio snapshot

To calculate the conditional variance of $R_D(t)$ at a given point in time t, we exploit the independence of the idiosyncratic risk factors $x_i(t)$ in the ASRF model, which implies that the Bernoulli variables $d_i(t)$ are independent of one another, conditional on ϵ_t, namely

$$
\begin{aligned}
\mathrm{Cov}_c(d_i(t), d_j(t)) &= \mathrm{Cov}\left(d_i(t), d_j(t) | \epsilon(t) = \epsilon_t\right) \\
&= \mathrm{E}_c\left(d_i(t) \cdot d_j(t)\right) - \mathrm{E}_c(d_i(t)) \cdot \mathrm{E}_c(d_j(t)) \\
&= P\left(x_i(t) \leq \frac{\Phi^{-1}(p) - \sqrt{\rho}\epsilon_t}{\sqrt{1-\rho}} \bigwedge x_j(t) \leq \frac{\Phi^{-1}(p) - \sqrt{\rho}\epsilon_t}{\sqrt{1-\rho}}\right) \\
&\quad - P\left(x_i(t) \leq \frac{\Phi^{-1}(p) - \sqrt{\rho}\epsilon_t}{\sqrt{1-\rho}}\right) \cdot P\left(x_j(t) \leq \frac{\Phi^{-1}(p) - \sqrt{\rho}\epsilon_t}{\sqrt{1-\rho}}\right) \\
&= 0
\end{aligned} \tag{11}
$$

for $i \neq j$. Given this fact and the identity of $E_c(R_D(t))$ and $E_c(d_i(t))$ from eq. (9) the conditional variance of the portfolio default rate can be derived as

$$
\begin{aligned}
\mathrm{Var}_c(R_D(t)) &= \mathrm{Var}_c\left(\frac{1}{N(t)} \cdot \sum_{i=1}^{N(t)} d_i(t)\right) \\
&= \frac{\mathrm{E}_c(R_D(t)) \cdot (1 - \mathrm{E}_c(R_D(t)))}{N(t)}.
\end{aligned} \tag{12}
$$

This expression is similar to the variance of default rates of uncorrelated borrowers. The difference is that in the latter case $E_c(R_D(t))$ can be simply replaced by the parameter p of the underlying Bernoulli distribution at any time t. The Wald confidence interval for the at confidence level $1-\alpha$, conditional on $\epsilon(t) = \epsilon_t$, is also of similar form as in eq. (3):

$$
R_{D,\alpha}^{u,l}(t) = r_D(t) \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\frac{r_D(t) \cdot (1 - r_D(t))}{N(t)}}. \tag{13}
$$

This is not surprising because, as discussed in section 2.3, in the ASRF model correlation between defaults exclusively stems from the coupling of all borrowers' asset values to the systematic factor.

3.1.2 Conditional default rate in the long-run average

Taking advantage of the results of section 3.1.1 we can perform the analogous calculation for the long-run average default rate R_L as defined in eq. (1). Its expectation value in the ASRF model, conditional on the realizations ϵ_t of the systemic risk factor in the pre-defined timeframe, reads

$$
\mathrm{E}_c(R_L) = \mathrm{E}_c\left(\frac{1}{T}\sum_{t=1}^{T} R_D(t)\right) = \frac{1}{T}\sum_{t=1}^{T} E_c(R_D(t)) \tag{14}
$$

where $E_c(R_D(t))$ is given in eq. (9). To obtain the conditional variance we use the conditional independence

$$\mathrm{Cov}_c(d_i(t_1), d_j(t_2)) = \mathrm{Cov}\left(d_i(t_1), d_j(t_2) \,|\, \epsilon(t_1) = \epsilon_{t_1},\, \epsilon(t_2) = \epsilon_{t_2}\right) = 0 \tag{15}$$

for $i \neq j$ or $t_1 \neq t_2$ (or both), which can be derived in analogy to eq. (11), and the result from eq. (12). The conditional variance of R_L then reads

$$\begin{aligned}\mathrm{Var}_c(R_L) = \mathrm{Var}_c\left(\frac{1}{T}\sum_{t=1}^{T} R_D(t)\right) &= \frac{1}{T^2}\sum_{t=1}^{T}\mathrm{Var}_c(R_D(t)) \\ &= \frac{1}{T^2}\sum_{t=1}^{T}\frac{\mathrm{E}_c(R_D(t))\cdot(1-\mathrm{E}_c(R_D(t)))}{N(t)}\,.\end{aligned} \tag{16}$$

This shows that, despite borrower defaults being implicitly correlated by the influence of the systematic factor, the variance of the long-run average default rate keeps a form similar to the well-known $\sigma^2 = \frac{p(1-p)}{N}$ which holds for uncorrelated defaults. The difference resides in the time dependence of $\mathrm{E}_c(R_D(t))$, which arises because of borrowers being subject to a systematic risk factor $\epsilon(t)$ incorporating effects of the economic cycle.

From eq. (16) we can derive the Wald interval at confidence level $1-\alpha$ for the through-the-cycle PD, whose estimator is R_L under the assumption of a fixed timframe:

$$R_{L,\alpha}^{u,l} = r_L \pm \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\cdot\frac{1}{T}\cdot\sqrt{\sum_{t=1}^{T}\frac{r_D(t)\cdot(1-r_D(t))}{N(t)}}. \tag{17}$$

If a bank decides to weight the annual default rates $R_D(t)$ unequally with weights w_t to calculate the long-run average, the expression for the confidence interval changes to

$$R_{L,\alpha}^{u,l} = r_L \pm \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\cdot\sqrt{\sum_{t=1}^{T} w_t^2\frac{r_D(t)\cdot(1-r_D(t))}{N(t)}}. \tag{18}$$

As an example, the result of eq. (17) is illustrated in Figure 1 as a 95% confidence interval ($\alpha = 0.05$) for R_L in a fictitious portfolio of 10000 borrowers. The solid line represents a time series of 10 observed annual default rates, with their own 95% Wald confidence intervals shown as error bars. The dashed lines are the center, the upper and lower bound of the confidence interval for R_L respectively. Obviously the uncertainty of the long-run average default rate is considerably smaller than the uncertainties of the default rates at times t.

3.2 Uncertainty in case of random choice of portfolio snapshots

In this subsection we consider portfolio snapshot dates t as randomly chosen to estimate a TTC PD which in this case must be understood as a hypothetical "all-time" average. Thus, in addition to the uncertainty from finite sample size at a given time t, the second type of uncertainty discussed in section 2.2 comes into play, namely uncertainty from

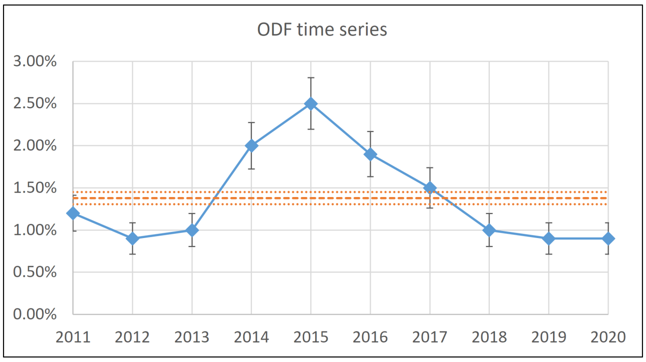

Figure 1: Example for a time series of $r_D(t)$ (solid line) with a 95% confidence interval for R_L due to limited sample size

the choice of t or, to put it differently, from the finite length of the timeframe $t = 1, .., T$. This uncertainty can only vanish in the limit $T \to \infty$, even in the limit $N \to \infty$. In a recent EBA staff paper [Casellina et al., 2021], this second type of uncertainty has been intensively studied in the limit $N \to \infty$, albeit with a focus rather on the tail of the distribution of default rates than on the long-run average. We will keep N finite in the following.

In contrast to the previous subsection where a confidence interval for the PD was derived from conditional variances, we now have to consider the *total variances* of default rates, see e.g. appendix B of [Gordy, 1998]. By the law of total variance we find

$$\mathrm{Var}(R_D(t)) = \mathrm{E}(\mathrm{Var}_c(R_D(t)) + \mathrm{Var}(\mathrm{E}_c(R_D(t)) \tag{19}$$

where $\mathrm{E}_c(R_D(t))$ and $\mathrm{Var}_c(R_D(t))$ have been calculated in the previous section. We hence have to calculate the expectation value of $\mathrm{Var}_c(R_D(t))$ and the variance of $\mathrm{E}_c(R_D(t))$ over all possible values of $\epsilon(t)$ given the standard-normal distribution of this risk factor. Assuming a constant portfolio size $N(t) = N$ this total variance is

$$\begin{aligned}\mathrm{Var}(R_D(t)) &= \mathrm{E}\left(\frac{\mathrm{E}_c(R_D(t)) \cdot (1 - \mathrm{E}_c(R_D(t)))}{N}\right) + \mathrm{E}(\mathrm{E}_c(R_D(t))^2) - \mathrm{E}(\mathrm{E}_c(R_D(t)))^2 \\ &= \frac{1}{N}\left(\mathrm{E}(\mathrm{E}_c(R_D(t)) - \mathrm{E}(\mathrm{E}_c(R_D(t))^2\right) + \mathrm{E}(\mathrm{E}_c(R_D(t))^2) - \mathrm{E}(\mathrm{E}_c(R_D(t)))^2,\end{aligned} \tag{20}$$

which can be simplified with the help of eq. (10) and a similar Gaussian integral calculation for the expectation value of $\mathrm{E}(\mathrm{E}_c(R_D(t))^2)$ to obtain the result

$$\mathrm{Var}(R_D(t)) = \frac{1}{N}\left(p - \Phi_2\left(\Phi^{-1}(p), \Phi^{-1}(p); \rho\right)\right) + \Phi_2\left(\Phi^{-1}(p), \Phi^{-1}(p); \rho\right) - p^2. \tag{21}$$

The function $\Phi_2(x, y; c)$ is the bivariate cumulative standard normal distribution for the variables x, y with the correlation parameter c. In the limit $\rho = 0$ the identity

$\Phi_2\left(\Phi^{-1}(p), \Phi^{-1}(p); 0\right) = p^2$ holds, hence the total variance reduces to the well-known form $\sigma^2 = \frac{p(1-p)}{N}$ for uncorrelated defaults.

Considering now the long-run average default rate (1), we find the total variance

$$
\begin{aligned}
\operatorname{Var}(R_L) &= \operatorname{Var}\left(\frac{1}{T}\sum_{t=1}^{T} R_D(t)\right) = \frac{1}{T^2}\sum_{t=1}^{T}\operatorname{Var}(R_D(t)) \\
&= \frac{1}{T\cdot N}\left(p - \Phi_2\left(\Phi^{-1}(p), \Phi^{-1}(p); \rho\right)\right) + \frac{\Phi_2\left(\Phi^{-1}(p), \Phi^{-1}(p); \rho\right) - p^2}{T}.
\end{aligned} \tag{22}
$$

Obviously the variance of R_L increases with increasing p and increasing ρ and decreases with increasing T and increasing N. In section 4 we will study this result numerically. The Wald confidence interval for the TTC PD under the assummption of a random choice of portfolio snapshots is [2]

$$
R^{u,l}_{L,\alpha} = r_L \pm \Phi^{-1}\left(1 - \frac{\alpha}{2}\right) \cdot \sqrt{\operatorname{Var}(R_L)}\Big|_{p=r_L}. \tag{23}
$$

with the expression for the variance from eq. (22) and the observed long-run average r_L as defined in eq. (2).

4 Discussion: Numerical evaluation and consequences for MoC

4.1 Numerical evaluation of Wald confidence intervals

In this section we evaluate our main analytical results, the Wald confidence intervals in eqs. (17) and (22), numerically to illustrate their size. This allows to compare the two possible approaches to each other and to the estimated PD level to which they refer. We assume the input parameters

- $\alpha = 0.05$, i.e. all confidence intervals are evaluated at 95% level, and
- $T = 10$.

From an empirical point of view, it has long been known from analyses of the empirical distribution of default rates measured by rating agencies that default rates exhibit large fluctuations. Research has shown, see e.g. [Schuermann and Hanson, 2004, Cantor et al., 2007], that for investment-grade default rates standard deviations of the same order of magnitude as the mean are not uncommon. Our results provide a theoretical underpinning of this empirical knowledge.

In Figure 2 the half width of the confidence interval defined by eq. (17) is shown for different portfolio sizes N and different long-run levels of default rates r_L, where the observation window which underlies r_L is treated as given. To keep this illustration

[2] In practical applications where the value of ρ for a particular portfolio is unknown, one might use the empirical variance of the observed annual values $r_D(t)$ instead of eq. (21) to estimate the uncertainty of R_L.

simple we neglect the variations of $r_D(t)$ over time, which are of minor importance is this case, and set $r_D(t) = r_L$. Note that the plot axes are logarithmically scaled.

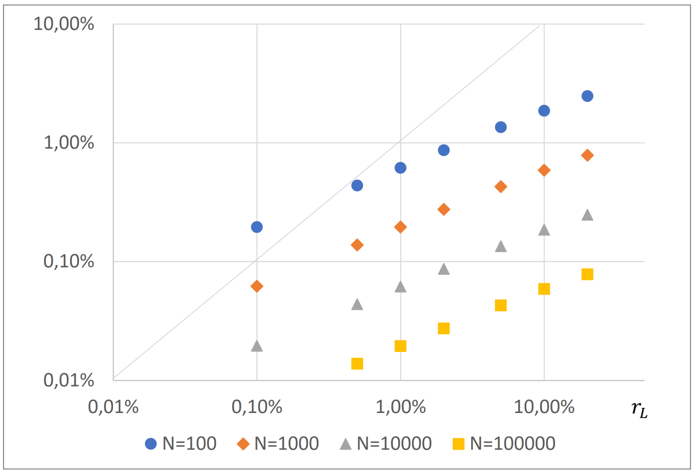

Figure 2: Uncertainty for a fixed timeframe: Half width of the 95% confidence interval for different values of r_L ($T = 10$).

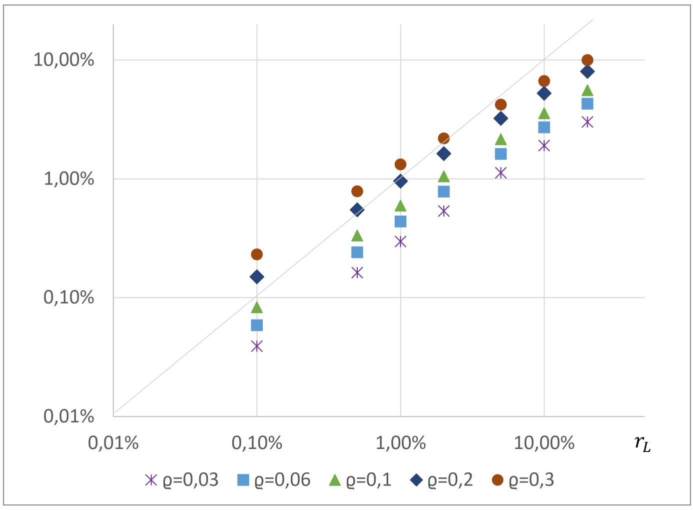

Figure 3: Uncertainty for a random choice of timeframe in the limit $N \to \infty$: Half width of the 95% confidence interval for different values of r_L ($T = 10$).

The figure demonstrates that statistical uncertainty of PD due to *limited sample size* can be of the same order of magnitude as the PD itself for small segments or rating grades with $N \sim 100$, in particular for PDs below 1%, i.e. for typical low-default portfolios. For larger segments, this will not be the case, as the uncertainty decreases with increasing N as expected. The expression from eq. (17) can readily be used to define a MoC of category C at a confidence level to be defined by the bank, assuming a given or prescribed choice of the observation window.

In Figure 3 the half width of the confidence interval defined by eq. (22) is shown for different values of the asset correlation ρ and different observed long-run levels of default rates r_L. In contrast to Figure 2 the observation window is treated as randomly selected. In order to separate the time-effect from the size-effect, we show the results here for the limit $N \to \infty$. Note that for $\rho = 0$ this confidence interval is of width 0. The figure demonstrates that, for the realistic value of 10 available annual historical time slices, statistical uncertainty of PD due to the *choice of the timeframe* can easily be larger than the uncertainty due to sample size (see above) if ρ is not assumed to be very small. It can also be of the same order of magnitude as the PD itself. This is the case although the ASRF model does not assume any autocorrelation in the time series of default rates. It is to be expected that autocorrelation would further increase the size of this type of uncertainty.

4.2 Numerical simulation of confidence intervals

To ensure that the approximations and shortcomings of a Wald confidence interval, see [Brown et al., 2001], do not lead to misleading conclusions we perform, in this subsection, a numerical simulation of the confidence interval of a TTC PD. To this end, we simulate defaults in a portfolio of N borrowers under the assumptions of the ASRF model, see section 2.3. We perform multi-year simulations with $T = 10$ for fixed values of N, p and ρ, whereby for each year ϵ and the x_i of the borrowers are randomly selected without any autocorrelation. After calculating the multi-year average of the resulting default rates over the simulation period, we repeat the procedure. This approach corresponds to the one treated analytically in section 3.2, i.e. to a random choice of the observation period.

After 1000 simulations the 95-th percentile of the 10-year average default rates is calculated and the predefined value of p is subtracted from it since it can be considered the "true" value of the TTC PD. The result corresponds to the half width of the confidence interval in an estimation of the TTC PD. We display the results for various values of p, N and ρ in figure 4. Note that the plot axes are logarithmically scaled.

As can be seen in this figure, the confidence intervals can be of the same order of magnitude as the target PD itself or even greater, depending on the amount of correlation among borrowers. This is the case even if the number of borrowers is large. We find that N does not play a significant role for large values of ρ where the uncertainty is completely dominated by the second type, i.e. the uncertainty stemming from the choice of the timeframe.

To illustrate how a Margin of Conservatism of the size indicated by these confidence intervals translates to capital requirements in the IRB approach, we show the effect of a PD increase, *ceteris paribus*, on IRB risk-weighted assets in table 1, using the formula in article 153 of the EU Capital Requirements Regulation and inserting a maturity of $M = 1$. As an example: For an estimated PD of 0.1% and a relative add-on of 100% due to statistical uncertainty, i.e. a final conservative PD of 0.2%, the capital requirement for credit risk is 61% higher than it would be without the add-on. Given the results of this chapter it becomes clear that MoC for statistical uncertainties in case of corre-

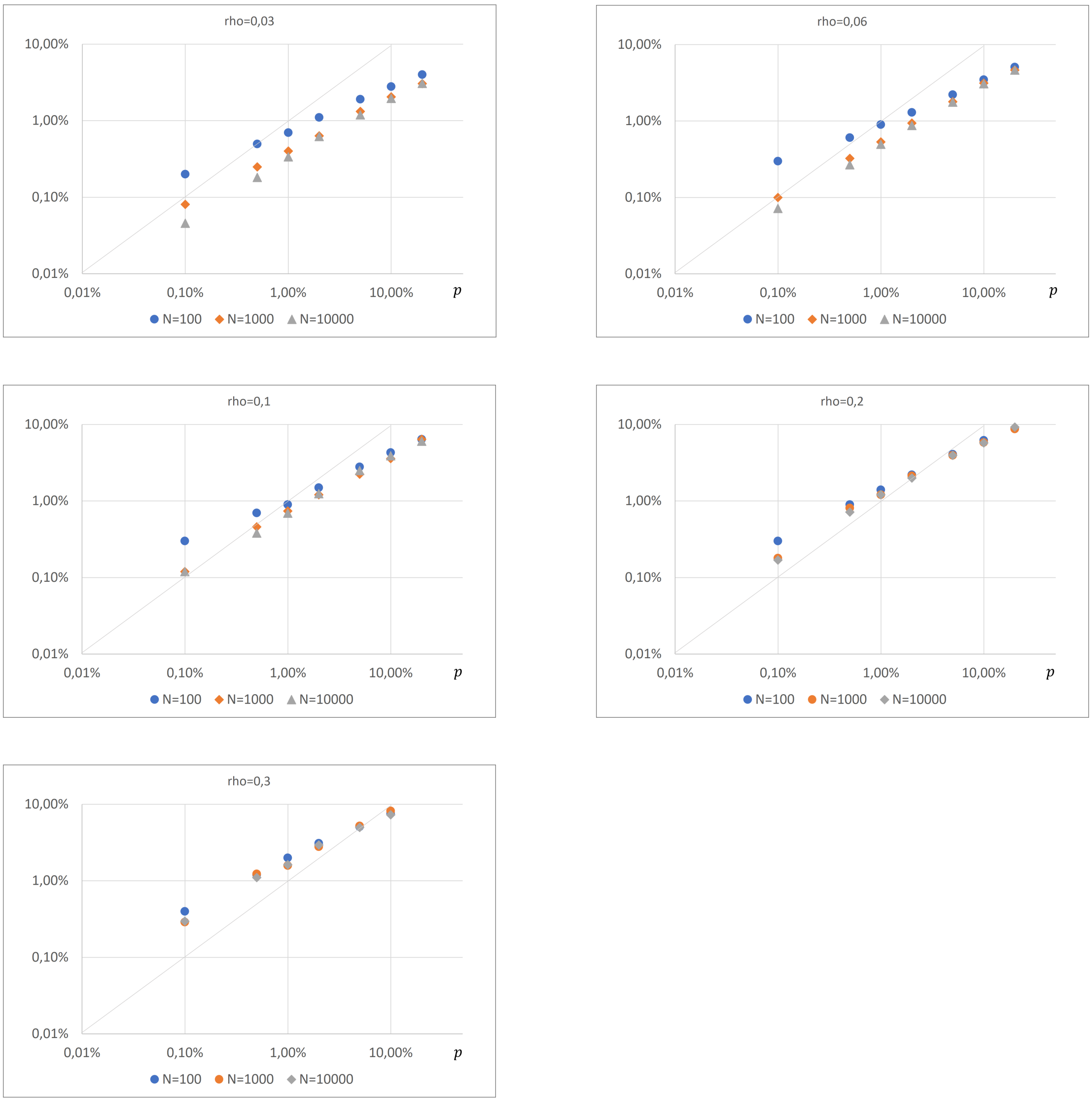

Figure 4: Half width of the simulated 95% confidence interval for different values of p, N and ρ ($T = 10$). The timeframe is considered to be randomly chosen.

lated borrowers and in particular the treatment of the timeframe as either predefined or randomly chosen has important consequences for banks' capital requirements.

PD	**+50%**	**+100%**	**+200%**
0.01%	1.39	1.75	2.41
0.10%	1.33	1.61	2.08
0.50%	1.23	1.40	1.66
1.00%	1.18	1.31	1.50
5.00%	1.18	1.33	1.55
1.00%	1.17	1.27	1.34

Table 1: Capital factors for different PD levels and a relative MoC add-on of 50%, 100% and 200% to the PD

5 Conclusions

We have shown in this article which two types of statistical uncertainties arise in PD estimation for a portfolio with correlated borrower defaults. We have shown

- how these uncertainties be quantified and combined in the ASRF model, i.e. the model underlying the risk weight formulas of the IRB approach and
- under which assumptions the choice of reference dates contributes to the uncertainty.

By means of a numerical evaluation and simulation we have demonstrated that a Margin of Conservatism of category C, if calculated as a 95% confidence interval around the PD under the ASRF model assumptions, may have important consequences for banks' capital requirements under the IRB approach. Important drivers of the size of statistical uncertainty are

- the confidence level,
- the size of the portfolio,
- the assumed or derived asset correlation between borrowers and
- the length of the historical timeframe.

A crucial debate is whether and how the choice of the historical timeframe underlying the estimation is to be included in the calculation of regulatory MoC. On the one hand banks are required to choose this timeframe following detailed regulations to reduce the risk that the observed average default rate of this timeframe differs materially from the "true" TTC PD. On the other hand, available data history is usually limited, often to one single macroeconomic cycle, so from a mathematical point of view the uncertainty from the choice of the timeframe cannot be fully excluded. As of now, IRB regulations leave room for interpretations in this point. An aspect which leaves room for future work is how, in case of calibration and MoC calculation on rating-grade level, the rating philosophy (PiT or TTC) influences the MoC. If the ratings exhibit a strong PiT behaviour, one expects the default rates of single rating grades to be rather constant in time, hence the uncertainty from the choice of the timeframe in determining the long-run average default rate of a grade is expected to be lower than in a TTC rating system, while it is not clear if the allocation of borrowers to grades is more uncertain in this case.

Acknowledgements: The authors are indebted to Marco Folpmers, Marcel Jegel, Heiner Möllers and Thomas Moosbrucker for valuable discussions and input to this work.

References

[BCBS, 2005] BCBS (2005). An explanatory note on the basel II IRB risk-weight functions.

[Brown et al., 2001] Brown, L. D., Cai, T. T., and DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical Science*, 16(2):101–133.

[Cantor et al., 2007] Cantor, R., Hamilton, D., and Tennant, J. (2007). Confidence intervals for corporate default rates. *SSRN Electronic Journal.*

[Casellina et al., 2021] Casellina, S., Landini, S., and Uberti, M. (2021). The estimation risk and the IRB supervisory formula. *EBA Staff Paper Series*, 11.

[EBA, 2017] EBA (2017). Guidelines on PD estimation, LGD estimation and the treatment of defaulted exposures (EBA/GL/2017/16).

[ECB, 2019] ECB (2019). ECB guide to internal models.

[Garcia-Cespedes and Moreno, 2016] Garcia-Cespedes, R. and Moreno, M. (2016). Probability of default uncertainty in the vasicek credit risk framework. *Available at SSRN: https://ssrn.com/abstract=2761013.*

[Gordy, 1998] Gordy, M. B. (1998). A comparative anatomy of credit risk models. *Board of Governors of the Federal Reserve System Research Series.*

[Miao and Gastwirth, 2004] Miao, W. and Gastwirth, J. (2004). The effect of dependence on confidence intervals for a population proportion. *The American Statistician*, 58:124–130.

[Schuermann and Hanson, 2004] Schuermann, T. and Hanson, S. (2004). Estimating probabilities of default. *Federal Reserve Bank of New York - Staff Report no. 190.*

[Vašiček, 2002] Vašiček, O. A. (2002). The distribution of loan portfolio value. *Risk.*