Calibration of Discrete Probability of Default Rating Scale

From Segment Default Averages to Grade-Level Estimates

Andrija Djurovic

www.linkedin.com/in/andrija-djurovic

Internal Ratings-Based Modeling Process

- The development of an Internal Ratings-Based (IRB) model consists of two main areas: the development of the risk differentiation function and risk quantification.
- Typically, these two steps are performed on different datasets, with significantly greater flexibility left to practitioners in the development of the risk differentiation function.
- Development of the risk differentiation function is the first step in the overall model development process. The main goal of this step is to ensure the model can adequately rank risk, providing meaningful differentiation between low- and high-risk obligors or facilities.
- After developing the risk differentiation function, practitioners perform the risk quantification process, which essentially involves assigning an appropriate risk level to the output of the risk differentiation function.
- In practice, there is typically a difference between the risk quantification process for continuous and discrete outputs of the risk differentiation function.
- Regardless of the type of output (continuous or discrete) of the risk differentiation function, risk quantification typically consists of specific steps such as: assignment of the base or best estimate, quantification of appropriate adjustments (AA), quantification of the margin of conservatism (MoC), and quantification of any other additional conservatism, for instance that required by certain local regulators.
- These steps apply to all three risk parameters: Probability of Default (PD), Loss Given Default (LGD), and Exposure at Default (EAD).

Regulatory Requirement for Probability of Default Model Calibration

- Based on the PD & LGD Guidelines, for the purpose of determining PD estimates in the calibration process, institutions should consider either:
  - the long-run average default rate (LRA DR) at the grade or pool level; or
  - the long-run average default rate at the calibration segment level.
- In the same document, the EBA provides an overview of combinations of rating and calibration philosophies and points out certain combinations that are not allowed for the purpose of own funds calculation.
- Furthermore, in its Guide to Internal Models (2025), the ECB clarifies in paragraph 241 its understanding that PD estimates without MoC should be equal to LRA DRs. Moreover, paragraphs 242 and 243 further emphasize potential situations of divergence between LRA DRs and raw PDs for two different calibration approaches.
- Although the above requirements represent only a small excerpt of the overall regulatory requirements on PD model calibration, they sufficiently cover the purpose of this presentation.
- This presentation focuses on the common statistical methods used to calibrate a discrete IRB PD rating scale, starting from the LRA DR at the calibration level and ending with grade-level PD estimates. As it does not explain the details of certain calibration requirements, practitioners are encouraged to link the presented methods to concrete requirements and critically adjust them to specific situations.

Probability of Default Model Calibration in Practice

One of the two presented calibration methods assumes that only the LRA DR at the calibration segment level is available for the purpose of rating scale calibration. Therefore, the only option practitioners have is to define initial or raw PD estimates, determine the most appropriate distribution of exposures across rating grades, and finally calibrate the rating grade PDs to match the LRA DR at the calibration segment level.

The three most common methods applied in practice for calibrating a rating scale in such situations are:

1. linear rescaling;
2. optimization of the intercept of log-odds-transformed raw PD estimates;
3. optimization of the intercept and slope of log-odds-transformed raw PD estimates.

The following subsections present the mathematical formulation of these three calibration methods.

Linear Rescaling

$$K = \frac{ct}{\frac{\sum_{i=1}^{G} n_i \, pd_i}{\sum_{i=1}^{G} n_i}}$$

where:

- ct denotes the central tendency, or LRA DR, at the calibration segment level;
- pd_i denotes the raw PD estimate for rating grade i;
- n_i denotes the number of exposures in rating grade i;
- G denotes the number of grades in the rating scale.

Probability of Default Model Calibration in Practice cont.

Intercept Optimization

$$\min_{a\in\mathbb{R}} \frac{\sum_{i=1}^{G} n_i \frac{\exp(a+lo_i)}{1+\exp(a+lo_i)}}{\sum_{i=1}^{G} n_i} - ct$$

Intercept and Slope Optimization

$$\min_{a,b\in\mathbb{R}} \frac{\sum_{i=1}^{G} n_i \frac{\exp(a+b\cdot lo_i)}{1+\exp(a+b\cdot lo_i)}}{\sum_{i=1}^{G} n_i} - ct$$

where:

- ct denotes the central tendency, or LRA DR, at the calibration segment level;
- lo_i denotes the raw PD estimate log-odds calculated as $\log\left(\frac{pd}{1-pd}\right)$;
- n_i denotes the number of exposures in rating grade i;
- G denotes the number of grades in the rating scale;
- a and b are the intercept and slope, respectively, subject to the optimisation problem.

Simulation Study

Assume the table below presents the PD rating scale that needs to be calibrated.

```
##         rating   no  nb   ng     pd
## 01 (-Inf,473)  413  63  350 15.25%
##  02 [473,501) 1269 134 1135 10.56%
##  03 [501,530) 1765  80 1685  4.53%
##  04 [530,556) 1163  28 1135  2.41%
##  05 [556,Inf)  845  10  835  1.18%
```

where:

- *rating* denotes the rating grade, with lower grades indicating riskier obligors;
- *no* denotes the number of exposures;
- *nb* denotes the number of observed defaulted exposures;
- *ng* denotes the number of observed non-defaulted exposures;
- *pd* denotes the raw PD estimate.

Simulation Goal

For the given raw PD estimates, which in this case equal the observed default rate ($\frac{nb}{no}$) from the development sample, the current segment PD is 5.77%.

The goal of this simulation is to calibrate the grade-level PD estimates to match the segment LRA DR of 7% using all three methods presented in the previous slides, and to compare the calibrated estimates graphically.

As this example serves only a demonstrative purpose, practitioners are encouraged to explore different calibration designs, including calibration to a lower segment LRA DR.

Simulation Results

Estimated Parameters

1. Linear rescaling (`K`): 1.212;
2. Intercept optimization (`a`): 0.213;
3. Intercept and slope optimization (`a` and `b`): 0.029, 0.928.

Calibrated PDs

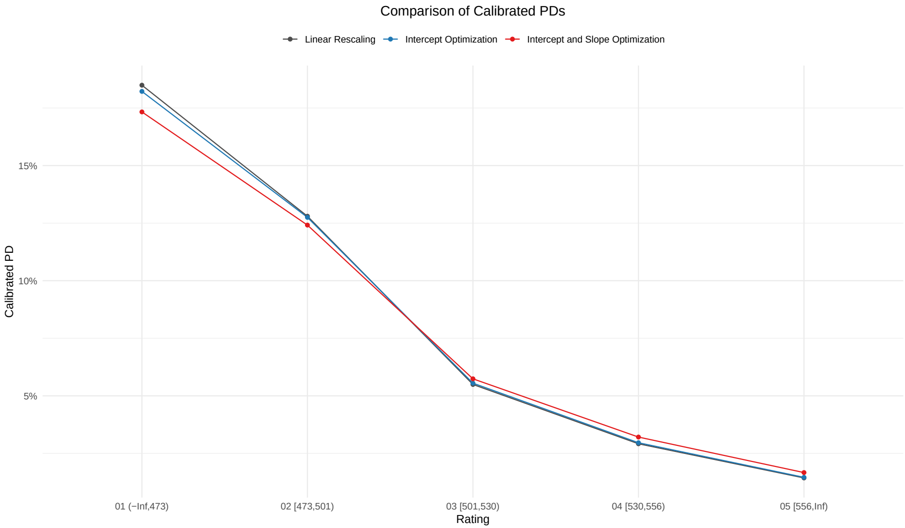
```
##         rating   no  nb   ng     pd pd.scaling pd.log.a pd.log.ab
##  01 (-Inf,473)  413  63  350 15.25%     18.49%   18.22%    17.33%
##   02 [473,501) 1269 134 1135 10.56%     12.80%   12.75%    12.42%
##   03 [501,530) 1765  80 1685  4.53%      5.49%    5.55%     5.74%
##   04 [530,556) 1163  28 1135  2.41%      2.92%    2.96%     3.21%
##   05 [556,Inf)  845  10  835  1.18%      1.43%    1.46%     1.67%
```

Segment PD Checks

$$\frac{\sum_{i=1}^{G} no_i \cdot \text{pd calibrated}_i}{\sum_{i=1}^{G} no_i}$$

1. Linear rescaling: 7.00%;
2. Intercept optimization: 7.00%;
3. Intercept and slope optimization: 7.00%.

Simulation Results cont.



Additional Notes

When implementing intercept and intercept and slope calibration methods, most practitioners define this as a specific optimization problem. In other words, the optimization problem is often defined as minimizing the difference between segment PD based on the calibrated estimates and the segment LRA DR. However, the same process for optimizing the intercept can be performed using the standard implementation of logistic regression. The following code snippet demonstrates this process in `R` and `Python`.

R Example

```r
#data import
url <- "https://raw.githubusercontent.com/andrija-djurovic/adsfcr/main/calibration/drs_calibration.csv"
rs <- read.csv(file = url,
               header = TRUE)
#segment central tendency
rs$ct <- 0.07
#regression model
mbc.a <- glm(formula = ct ~ 1,
             family = "binomial",
             weights = rs$no,
             offset = log(rs$pd / (1 - rs$pd)),
             data = rs)
#optimized intercept
coef(mbc.a)
```

```
## (Intercept)
##   0.2134049
```

```r
#calibrated pds
pd.mbc.a <- predict(object = mbc.a, type = "response")
pd.mbc.a
```

```
##          1          2          3          4          5
## 0.18221778 0.12751147 0.05550961 0.02963321 0.01460841
```

```r
#check - segment pd
sum(pd.mbc.a * rs$no / sum(rs$no))
```

```
## [1] 0.07
```

Additional Notes cont.

Python Example

```
import pandas as pd
import numpy as np
import statsmodels.api as sm
import statsmodels.formula.api as smf
#data import
url = "https://raw.githubusercontent.com/andrija-djurovic/adsfcr/main/calibration/drs_calibration.csv"
rs = pd.read_csv(filepath_or_buffer = url)
#segment central tendency
rs["ct"] = 0.07
#regression model
mbc_a = smf.glm(formula = "ct ~ 1",
                data = rs,
                family = sm.families.Binomial(),
                freq_weights = rs["no"],
                offset = np.log(rs["pd"] / (1 - rs["pd"]))).fit()
#optimized intercept
mbc_a.params
```

```
## Intercept    0.213405
## dtype: float64
```

```
#calibrated pds
pd_mbc_a = mbc_a.predict()
pd_mbc_a
```

```
## array([0.18221778, 0.12751147, 0.05550961, 0.02963321, 0.01460841])
```

```
#check - segment pd
np.sum(pd_mbc_a * rs["no"]) / np.sum(rs["no"])
```

```
## np.float64(0.07)
```