Estimating Probabilities of Default for Low Default Portfolios

Pluto-Tasche Approach

Andrija Djurovic

www.linkedin.com/in/andrija-djurovic

Low Default Portfolio

- Qualitative descriptions of an Low Default Portfolio (LDP) leave ample room for interpretation, as different individuals may have varying opinions on whether a given portfolio qualifies as an LDP.
- The low number of defaults within a LDP undermines estimates' reliability and statistical validity for quantitative risk parameters based on historical default experience.
- LDPs are not necessarily low-data portfolios. The scarcity of defaults needs to be considered in relation to the size of the portfolio producing them.
- Regulators may be concerned that Probability of Default (PD) estimates based solely on simple historical averages or judgmental considerations may underestimate the bank's capital requirements due to default scarcity.
- The Pluto-Tasche (PT) approach is one way to address the estimation of PD for the LDP.
- Alternative methods are also available and warrant examination and comparison with the PT approach. For details, practitioners can refer to the adjustment of the Alan Forrest and the BCR methods.

Pluto-Tasche Proposal for the PDs Estimation

The following steps outline the data preparation process for running the PT PD estimation for the multi-year design at the grade level, which accounts for asset correlation and year-to-year correlation:

1. Identify the LDP.
2. Define the historical sample for the LDP (multi-year period) using internal data.
3. The sample should specify each obligor's grade at the start of each year within the historical period.
4. Aggregate multi-year data by obligor rating, including the number of obligors and defaults.
5. Select the algorithm inputs: asset correlation, year-to-year correlation, confidence level, and number of simulations.
6. Run the PT estimation process for each rating grade.
7. (Optional) Rescale the estimated PDs to match the target portfolio's central tendency.

More details on the method and data preparation are available here.

The following slides briefly describe the Monte Carlo process used for PD estimation and present results from the simulation study based on the same dataset as the other two methods mentioned in the previous slide. For comparison, practitioners can refer to these presentations: Adjustment of the Alan Forrest and BCR.

Pluto-Tasche Method - Upper Bound of the Portfolio PD

1. The change in the value of obligor i's assets over a year t is given by:

$$y_{i,t} = \sqrt{\rho} z_t + \sqrt{1-\rho}\epsilon_{i,t}$$

where ρ is asset correlation, z and ϵ_i systemic and idiosyncratic factor, respectively.
The obligor defaults if the asset value is lower than the value of c for the specific year. Given the above assumptions, we are interested in the probability (PD):

$$P(y_{i,t} \leq c) = PD_{i,t} = PD$$

which describes the long-term average 1-year probability of default among the obligors that have not defaulted before and may be considered a through-the-cycle PD.

2. Given the number of years (T) and year-to-year correlation (θ) we simulate the systemic factor for each year as follows:

$$z_{t,t \leq T} = \theta z_{t-1} + \sqrt{1-\theta^2}\epsilon_t$$

where $z_{t,t=1}$ and ϵ_t are drawn from the standard normal distribution (`N[0, 1]`).

3. Using the $z_{t,t \leq T}$ we define the conditional PD (PD_{c_t}) for each year $t, t \leq T$ as follows:

$$PD_{c_t} = N\left[\frac{N^{-1}(PD) - z_t\sqrt{\rho}}{\sqrt{1-\rho}}\right]$$

where N and N^{-1} represent the distribution and quantile function of the normal standard distribution, respectively.

Pluto-Tasche Method - Upper Bound of the Portfolio PD cont.

4. For the multi-year period, we calculate the cumulative PD as follows:

$$PD = 1 - \prod_{t=1}^{T} (1 - PD_{c_t})$$

5. Given the total number of obligors n up to the selected rating, the total number of defaults k also up to the selected rating, and the cumulative probability of default PD, the likelihood of observing no more than k defaults among n obligors is calculated as:

$$1 - \gamma = \sum_{i=0}^{k} \binom{n}{i} PD^i (1 - PD)^{n-i}$$

   where γ denotes the selected confidence level.

6. Finally, we determine the upper bound of the rating PD through numerical optimization for PD using Monte Carlo simulations to minimize the difference between the average simulated and the chosen confidence level (γ).

Simulation Study - R Code

```r
#source r script
source("https://raw.githubusercontent.com/andrija-djurovic/adsfcr/main/ldp/pt.R")

#inputs
n <- c(26, 122, 182, 123, 24, 14, 9)       #number of obligors per rating
k <- c(0, 0, 0, 0, 1, 1, 2)                #number of defaults per rating
theta <- 0.30                              #year-to-year correlation
rho <- 0.12                                #asset correlation
T <- 5                                     #number of years
cl <- 0.75                                 #confidence level
N <- 1e4                                   #number of simulations

#pluto-tasche pd estimates
pd <- ldp.pt(n = n,
             k = k,
             theta = theta,
             rho = rho,
             T = T,
             cl = cl,
             N = N,
             seed = 123)

#upper bound of the pd
sprintf("%.2f%%", 100 * pd)
```

```
## [1] "0.38%"  "0.40%"  "0.53%"  "1.03%"  "3.43%"  "5.51%"  "10.44%"
```

Simulation Study - Python Code

```python
#source python script
import requests
url = "https://raw.githubusercontent.com/andrija-djurovic/adsfcr/main/ldp/pt.py"
r = requests.get(url)
exec(r.text)

#inputs
n = [26, 122, 182, 123, 24, 14, 9]     #number of obligors per rating
k = [0, 0, 0, 0, 1, 1, 2]              #number of defaults per rating
theta = 0.30                           #year-to-year correlation
rho = 0.12                             #asset correlation
T = 5                                  #number of years
cl = 0.75                              #confidence level
N = int(1e4)                           #number of simulations

#pluto-tasche pd estimates
pd = ldp_pt(n = n,
            k = k,
            theta = theta,
            rho = rho,
            T = T,
            cl = cl,
            N = N,
            seed = 123)

#upper bound of the pd
[f"{round(100 * value, 2):.2f}%" for value in pd]
```

```
## ['0.38%', '0.40%', '0.53%', '1.03%', '3.45%', '5.50%', '10.46%']
```