Funds Peer Benchmarking Report

Appendix: Database & Methodology

Version: 1.0

Last modified: 15/05/26

Author: Artur Stypulkowski

Global Credit Data
by banks for banks

Contents

1. Purpose of the appendix

This appendix accompanies the public Funds Peer Benchmarking Report ("Report") and provides supporting information on the data sources, reference dataset, filtering approach, and definitions applied in the analysis. It is intended to help readers understand the scope of the underlying data and the methodological choices used to prepare the results presented in the report.

2. Overview of data sources

2.1. Special Fund Collection

We conducted a bespoke data collection for this Report, with contributions from 13 banks. Participant list is available in the Report. The dataset covers observations from 2007 to 2025 and includes more than 40,000 distinct funds. It provides information on portfolio composition, ratings, probabilities of default, observed defaults, recoveries, collateral, and selected structural characteristics relevant for credit risk analysis.

The report focuses on the main fund types represented in the dataset: mutual funds, hedge funds, other funds, private equity funds, pension funds, fund of funds, private debt funds, and real estate investment trusts.

The underlying collection also includes additional fund types not shown separately in the report, including infrastructure funds, money market funds, natural resources funds, sovereign wealth funds, and venture capital funds.

2.2. GCD LGD and EAD Data Pool

The Report also includes analysis drawn from the fund-related segment of GCD's LGD and EAD Data Pool. The broader database is the world's largest non-retail bank loan loss database, containing around 340,000 defaulted facility observations representing non-retail Basel asset classes. Built from submissions by more than 60 banks over a period of 20 years, the database provides detailed information on defaults, cash flows, collateral, and realized loss experience. For this Report, the relevant subset was used to analyze loss and recovery outcomes for fund exposures (see 3.3 for specific filter criteria). LGD data presented in the Report is separate from Special Fund Collection and contains data submitted by different banking institutions.

3. Reference dataset and filtering approach

3.1. Reference dataset

GCD provides members with the full "raw" dataset after the validation and auditing process. The

data is not filtered, cleansed or aggregated. GCD advises banks to create a reference dataset (RDS) from the full dataset: a subset of observations (borrowers, loans, collateral) that should resemble the internal target portfolio.

The RDS can be used for credit risk modelling, portfolio benchmarking, and the validation or calibration of models for a target portfolio. To be effective, the RDS should have two key qualities: representativeness and data quality. Creating a well-defined RDS is a critical success factor in the effective use of pooled data.

The full GCD data consists of data contributed by banks that have chosen to become GCD members. These banks' geographical lending footprint, loan and borrower types, lending and recovery practices, are combined in the database. Given its size, long time series and contributions from many countries, the dataset may be considered as broadly representative of an average bank. More precisely, however, it represents the average of GCD member banks, weighted towards the largest members which contribute the most data.

Furthermore, the GCD Representativeness Focus Group has collaboratively developed a comprehensive set of guidelines for assessing representativeness. The resulting document, available on our website, provides detailed guidance on methodologies, best practices, and key considerations for representativeness assessments. It also includes specific recommendations for the effective use of GCD's pooled datasets.

3.2. Filters applied to Special Fund Collection

POSITIVE OUTSTANDING AMOUNT

The Report includes only records for investment fund entities with positive outstanding amounts at the time of rating assignment (reporting date).

FURTHER FILTERING

The dataset was subject to data quality validation checks, and records that fell outside expected parameter ranges were excluded from the analysis. For example, observations with risk weights exceeding 1250% were removed to ensure consistency with regulatory and methodological standards. Where possible, these exclusions were reviewed and aligned with the submitting banks to ensure that records were excluded appropriately and consistently.

3.3. Filters applied to LGD & EAD data pool

REFERENCE DATASET (RDS)

The RDS filter, further explained in the appendix to GCD's Recovery Rates Reports, has been applied. This is a standard filter used in all LGD analyses and includes controls related to

data quality as well as adjustments addressing specific characteristics of pooled data.

ONE MILLION EUR FACILITY OUTSTANDING

Facilities with relatively small outstanding at the time of default have been filtered out to ensure only material default events are presented.

FURTHER FILTERING

The dataset was subject to data quality validation checks, and records that fell outside expected parameter ranges or were assessed as not representing fund exposures were excluded from the analysis. Where possible, these exclusions were reviewed and aligned with the submitting banks to ensure that records were excluded appropriately and consistently.

4. Definitions and methodological notes

DEFAULTED OBLIGORS

The Basel definition of default is used. According to the GCD Data Pool Regulations, all Pool participants must report their entire portfolio per asset class, including all the observed defaults. GCD verifies this requirement during the audit phase of each submission cycle.

COHORT SIZE

For default rate calculation, cohort size is defined as the number of performing obligors on 1st January of each year. In the dataset, this is equivalent to the number of obligors in the fourth-quarter snapshot of the previous year.

The number of obligors may change during the year as new obligors enter the portfolio and existing obligors leave. However, only the number of performing obligors at the start of the year is used to calculate the default rate. This number is reported as the cohort size.

When cohort sizes are reported over multiple years, the annual cohort sizes, as defined above, are averaged.

OBSERVED DEFAULT RATES

The default rate is calculated as the ratio of defaulting obligors during the period of interest to the number of performing obligors at the start of that period. In this report, the period of interest is always a one-year horizon starting on 1 January. Therefore, the denominator is the number of performing obligors at the start of the year in the portfolio of interest.

A performing obligor is an obligor whose internal rating is not defaulted on a specific

reporting date.

For a given year, the default rate is calculated using the fixed cohort of performing obligors at the start of the period, without considering a rolling calculation across all quarters.

Defaults are included in the numerator only if they relate to obligors in this starting cohort.

$$Yearly\ Observed\ Default\ Rate = \frac{\text{Defaults orruring withing the year}}{\text{Performing borrowers at the start of the year}}$$

RATING GRADE / INTERNAL RATINGS

GCD collects bank-internal information from participating banks. This gives members a comprehensive view of bank lending portfolios, including funds that do not have external ratings.

Each member estimates through-the-cycle PDs and internal ratings. To enable comparability across internal rating scales, members map their internal ratings to S&P equivalents, in line with their own regulatory external reporting.

REGION

The GCD dataset includes country-level information, based on the country of residence. GCD returns this information at country level in the dataset. For this report, country-level information has been aggregated to regional level.

The regions used in this report are:

- Africa
- Middle East
- Asia
- Oceania
- Europe
- America

IRB RISK WEIGHT

The IRB Risk Weight is calculated using the bank's internal estimates of Probability of Default (PD) and Loss Given Default (LGD), together with the official risk-weight calculation formula. Where an obligor has multiple facilities with different LGD levels, the risk weight is calculated as an exposure-weighted average and reported at obligor level. Averages displayed in the report are arithmetic averages unless otherwise stated.

LOSS GIVEN DEFAULT (LGD)

Global Credit Data (GCD) members do not provide LGD as an input field to the database but the underlying raw information such as outstanding amount at default or cash flows during the default. GCD calculates LGD by following different methodologies. Internal methodology as well as regional regulatory requirements might result in different calculation methods which member banks are able to realize on the raw data they receive back. In this report the following LGD definition is used:

$$LGD = 1 - \frac{\text{Recoveries} - \text{Costs}}{\text{Outstanding Amount at Default} + \text{Principal Advance} + \text{Cash out on Contingent L.} + \text{Financial Claim}}$$

Recoveries include
TT100 – Principal Payments
TT200 – Interest Payments
TT250 – Recorded Book Value
TT490 – Fees & Commissions Received
Outstanding Amount at Resolution

Costs include
TT500 – Legal Expenses
TT600 – Administrator/Receiver Fees
TT700 – Liquidation Expenses
TT800 – Other External Workout Costs
(No indirect costs)

Other Transaction Types
TT400 – Principal Advance
TT410 – Cash Out on Contingent Liability
TT420 – Financial Claim
TT450 – Interest Charged
TT480 – Fees & Commissions Charged

Cash flows are discounted by 3-month EURIBOR plus 5%. Post default drawings are treated as part of the Default Amount. All LGDs are floored at -50% and capped at 100%. The LGDs are calculated on facility level.

TTC PD

GCD collects the through-the-cycle PDs estimated by the lenders using their internal models.

FUND TYPES

The table below summarizes the fund types collected for the analysis and provides a description of each category. For some more granular fund types, the Special Fund Collection includes PD and risk weight data only. LGD data was not collected for these fund types.

Fund type	**Description**
Mutual fund	A collective investment vehicle managed by a professional fund manager or bank that pools investor capital and invests in specific asset classes. This category includes ETFs and unit trusts and may be geared or ungeared.
Hedge fund	An investment fund that typically uses derivatives, options, and other non-traditional strategies to take directional positions in asset markets.
Pension fund (provident, retirement, etc.)	A pooled investment vehicle established to manage savings for retirement and related long-term benefit obligations.

Private equity fund	A fund that invests directly in private companies or takes public companies private, with the aim of enhancing value over time and exiting through an initial public offering or sale.
Private debt fund	A fund that provides non-bank lending, including direct lending and investments in distressed debt.
Fund of funds	A fund that invests in other funds rather than directly in underlying assets.
Other fund	Other or unknown fund type.
Venture capital fund	A fund that invests in early-stage companies with high growth potential.
Infrastructure fund	A fund that invests in infrastructure assets such as transport, utilities, and energy.
Natural resources fund	A fund that invests in assets or companies related to energy, mining, agriculture, and similar natural resource sectors.
Money market fund	A fund that invests in short-term, high-quality debt instruments and aims to preserve capital while maintaining liquidity.
Sovereign wealth fund	A state-owned investment fund that manages national savings or foreign currency reserves with the aim of generating long-term returns.
Real estate investment trust (REIT)	A specialized type of fund investing in real estate. It may be geared or ungeared, listed or unlisted, and may hold single-property or multi-property portfolios.

About Global Credit Data

At GCD, we pool credit loss data directly from banks' books, providing vital insights into the financial industry since 2004.

As a non-profit organization owned by over 50 member banks, we focus on collecting detailed credit risk data, particularly for low-default portfolios.

Beyond data pooling, we offer a platform for knowledge exchange and research. We are actively engaged in understanding and assessing climate risk, demonstrating our commitment to addressing contemporary and future financial challenges.

Joining GCD gives you access to an exclusive community of banks and deep data insights, helping you gain market understanding and benchmark your performance.

For additional information, please contact

Olivier Plaetevoet
PD & Rating Executive
olivier.plaetevoet@globalcreditdata.org

Artur Stypulkowski
Senior Quantitative Analyst
artur.stypulkowski@globalcreditdata.org

Global Credit Data
by banks for banks

www.globalcreditdata.org