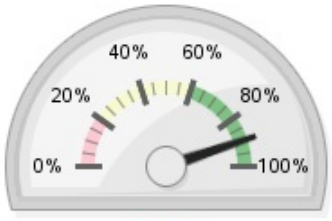
Global LGD Data Quality Dashboard | H1 2020

Global Credit Data
by banks for banks

GCD Total Data Quality Score

The GCD loan level data to model loan losses contains +230k facilities. The total Data Quality score is a stable 90% and current efforts target the improvement on Completeness of Optional Fields.

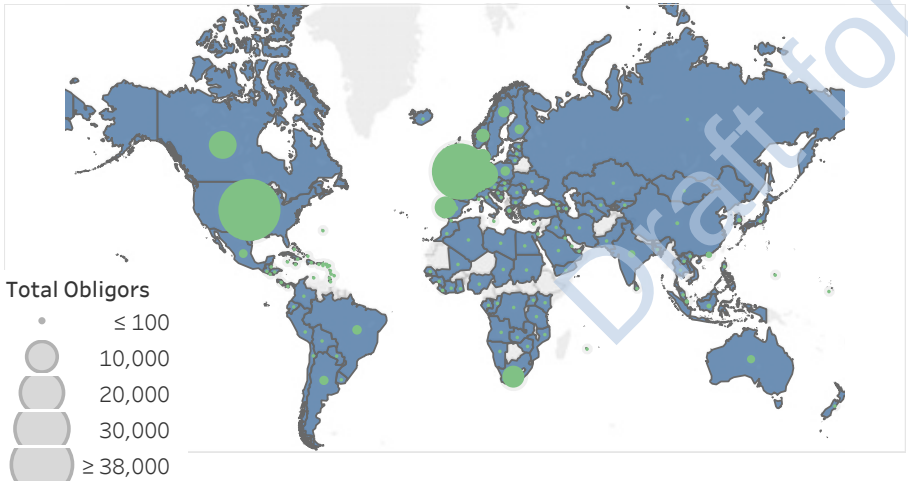


90%

Summary

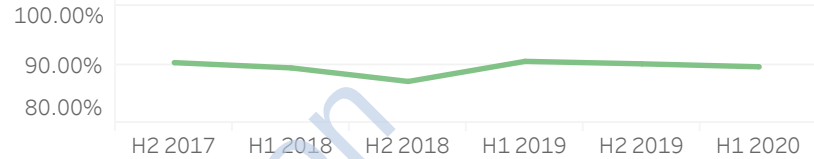
- The total Data Quality Score is based on the latest H1 2020 data submission cycle.
- The minimal decrease compared with the previous submission can be considered as a normal variation of data quality.
- The GCD Total Data Quality Score is calculated as the weighted average of the 5 Data Quality Dimensions listed to the right.

LGD Data Concentration



This chart shows how the +230k facilities are distributed world wide by 53 active member banks, including all 11 Facility Asset Classes. See the next page for an overview by Asset Class.

GCD Total Data Quality Score Trend



The new scoring was initiated in 2017 and members have 3 years to adjust to cyclically improved standards and requirements and resubmit the data accordingly.

Data Quality Dimensions (BCBS239 Compliant)

Scoring Type	# Checks	H1 2019	H2 2019	H1 2020
Comparability	6	96%	96%	96%
Stability	8	97%	97%	96%
Completeness of Optional Fields	98	68%	68%	66%
Consolidated Correctness	27	98%	98%	98%
Timeliness	1	91%	89%	88%

The methodology to calculate the scores for each data quality dimension is described in the GCD Data Quality Management Policy - appendix. The metrics for the different dimensions follow different concepts. Whereas the Comparability metrics is defined on a functional basis for the key risk drivers for LGD modelling purpose, the Correctness checks are tecnical and consolidated from 400 mandatory field level checks to 27 check categories. The metrics cover the central data quality dimension requirements from BCBS239.

Freshness of 2020 Data

Total members active in the LGD Pool	53
Full Submissions Performed	35
Full Submissions in part of Asset Classes	4
No full submissions Performed	10
Members onboarding	3
Member Data Submissions - Not Accepted	1

A full submission concerns a submission of all historical defaults from a member in all Facility Asset Classes a member is contributing to. The numbers reflects the refresh of default data over the last three years.

Draft for Validation

Copyright 2020 The Global Credit Data Consortium all rights reserved

Global LGD Datapool Overview | H1 2020

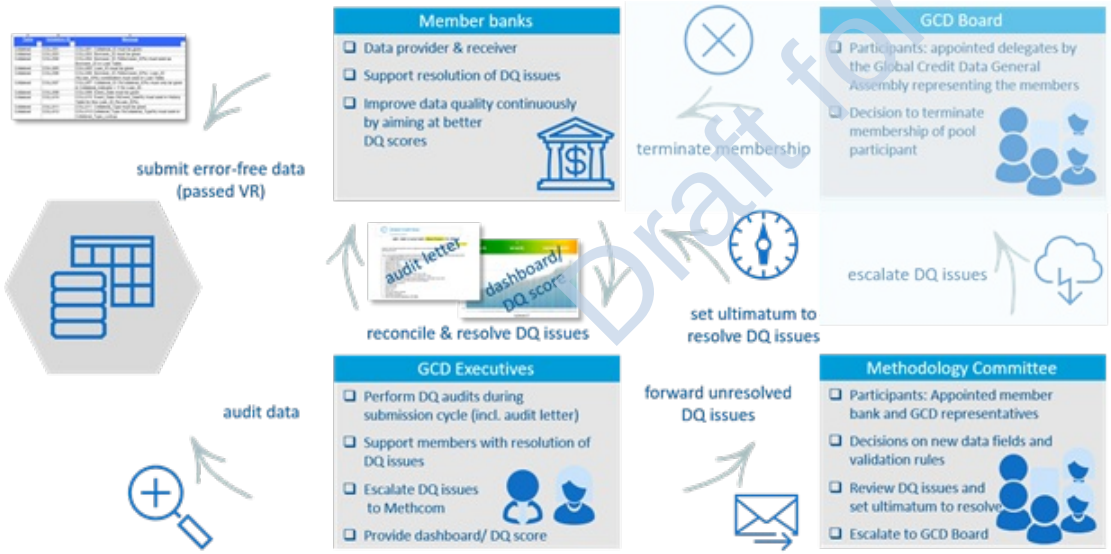
LGD H1 2020 Datapool Overview by Facility Asset Class

Facility Asset Class	Exposure (in mn EUR)	Number of Facilities	Number of Obligors
Small/Medium Enterprises (SME)	125,822	152,229	85,531
Large Corporates	282,325	37,372	17,656
Real Estate Finance	76,854	25,297	15,911
Private Banking	8,635	8,813	5,479
Banks & Financial Companies	87,505	4,394	2,723
Ship Finance	15,409	1,492	806
Project Finance	13,864	921	452
Commodities Finance	8,793	672	393
Aircraft Finance	5,248	709	320
Public Services	836	253	161
Sovereigns, Central Banks	11,569	187	122
Grand Total	636,861	232,339	129,554

Global Credit Data
by banks for banks

Datapool Volumetrics

GCD has built up the world's largest non-retail bank loan loss database with over 200,000 defaulted facility observations totalling over €500 billion in all non-retail Basel asset classes. The Global Credit Data LGD/EAD platform, the largest of GCD's databases, was created in 2005 and has grown rapidly ever since. It now gathers detailed information on tens of thousands of defaulted counterparties, particularly data on cash flows and collateral.



GCD Data Quality Management

To ensure data quality, GCD established and implemented an effective data quality management framework from the very beginning in 2005. The framework evolved since then to achieve compliance with leading regulatory standards like, e.g., BCBS 239 or ECB guide to internal models. The data quality governance framework is depicted in the figure, it shows all involved stakeholders, their interactions and roles.
The prevailing goal of the framework is to continuously measure, control and improve data quality. This allows members to broadly use GCD data and to be compliant with regulatory requirements.

Draft for Validation

Copyright 2020 The Global Credit Data Consortium all rights reserved

Global LGD Datapool Background | H1 2020

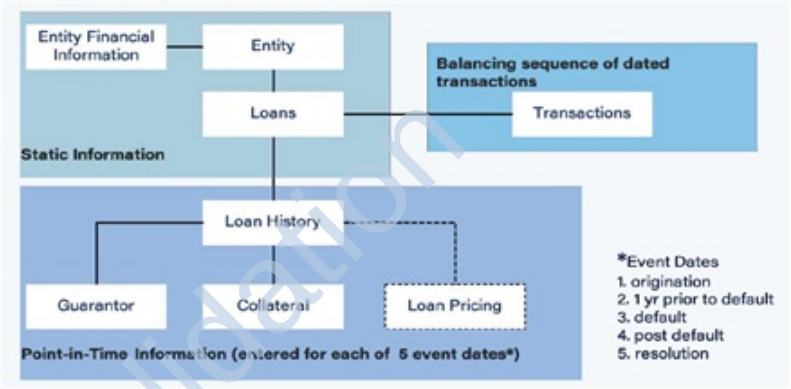
Global Credit Data
by banks for banks

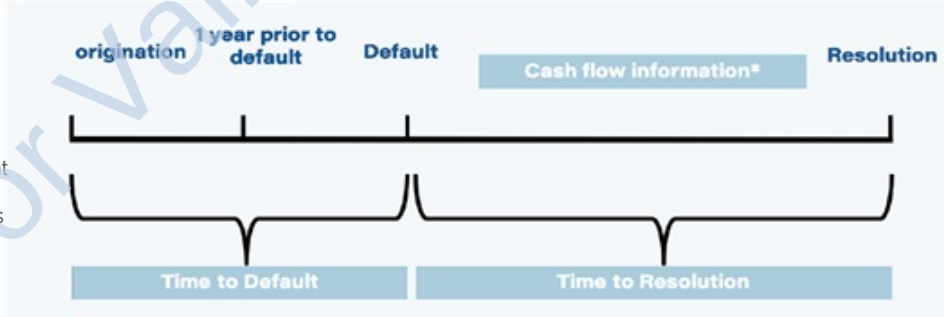
TABLES AND FIELDS

The structure of the database reflects the full complexity of the legal relationship between a bank lender and a borrower. It is designed to deal with the simplest through to the most complex deals. The types of complexity covered include:

- Single company borrower with multiple facilities (loans, commitments and offbalance sheet)
- Each loan or group of loans having single or multiple full or partial guarantors
- Each loan or group of loans being fully or partially secured by multiple collaterals

In addition, the information around loans, guarantors, collateral and pricing is time stamped so that the changing loan limits, collateral values and guarantor coverage can be reflected at different points in the life of the loan from origination through to resolution. The data that is used in the LGD/EAD database is recognised as the industry standard and is used by many banks for their own internal data collection to build a predictive model of lending for the future

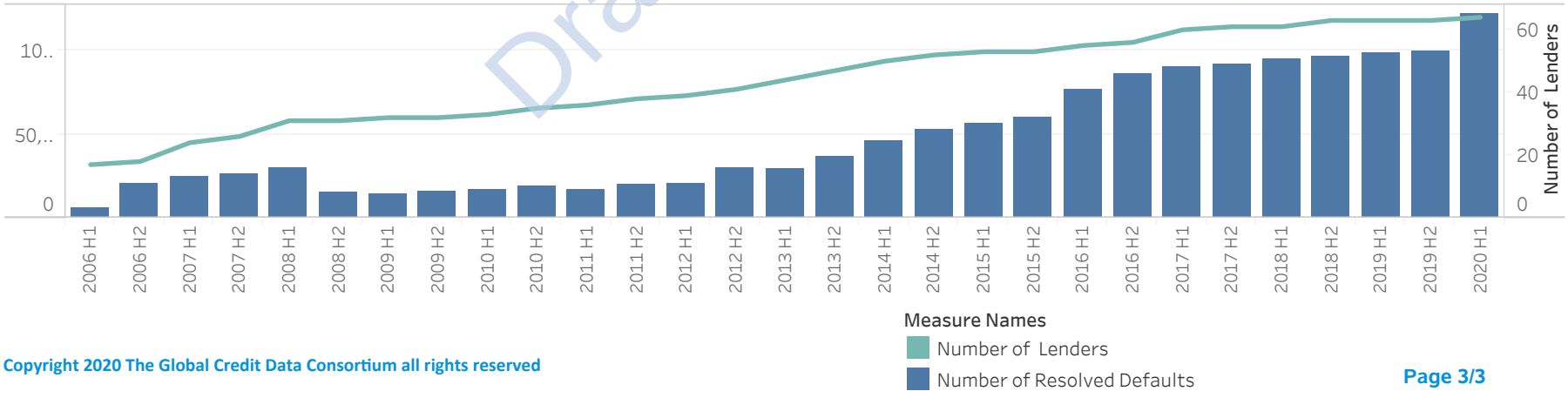




THE LGD/EAD DATABASE STRUCTURE

Data is collected from the point of origination, 1 year prior to the default, the point of the default and later time stamped data (unresolved cases only) up to and including the point of resolution (resolved cases only). Cash flow and accrual transactions are collected in relation to loans, collateral and guarantors and are time stamped to exact dates, sources and purposes. This is key for discounting of cash flow to produce economic LGDs and for calculating the progress of defaulted loan cases over time.

LGD Datapool Historical Growth



Draft for Validation

Copyright 2020 The Global Credit Data Consortium all rights reserved