



LGD Report 2019 - Large Corporate Borrowers

Confirmed: banks recover 76% from Large Corporate defaults

TABLE OF CONTENTS

SUMMARY	1
INTRODUCTION	2
REFERENCE DATA SET	2
ANALYTICS	3
CONCLUSIONS	7
APPENDIX: DATABASE & METHODOLOGY	

ABOUT GCD

Global Credit Data (GCD) is a non-profit association owned by 50+ member banks with the simple mission to help banks better understand and measure their credit risks through data pooling and benchmarking activities.

GCD started collecting historical loss data in 2004, to which member banks have exclusive access. This database now totals over 185,000 non-retail defaulted loan facilities from around the world.

In 2009 GCD introduced a PD database which now has over 15 years of default rates and PDs. GCD also runs a name and cluster benchmarking database to help banks calibrate and benchmark their PD, LGD and EAD models.

GCD operates all databases on a “give to get” basis, meaning that members must supply high quality data to receive data in return. The robustness of GCD’s data collection infrastructure place our databases as the global standard for credit risk data pooling.

CONTACT GCD

Nunzia Rainone
Analyst & Member Support Executive
nunzia.rainone@globalcreditdata.org

Nina Brumma
Head of Analytics and Research
nina.brumma@globalcreditdata.org

www.globalcreditdata.org

SUMMARY

- This is the GCD annual Report on Loss Given Default (LGD) for Large Corporate, in which numerical evidence of recoveries and losses is presented. The data set covers Large Corporate (>€50m turnover) borrowers who are recorded as defaulted in bank loan books, using the Basel default definition.
- GCD’s data pools support the key parameters of banks’ credit risk modelling (PD, LGD, EAD). This report covers LGD and represents a unique resource for all types of credit risk modelling: regulatory capital; pricing; stress testing; or expected loss provisioning models.
- The GCD data is always growing, from new member banks, and more years of default included. As expected, the results are consistent over time, confirming data stability and reliability.
- The results in this study offer an overall insight into the data on a global level and confirm the drivers, their direction and their levels shown in the 2018 report. The main findings are:
 - Seniority and collateral are confirmed as LGD drivers (26% senior unsecured vs 38% subordinated unsecured at obligor level. The total secured LGD is 22%).
 - LGD varies over time, indicating that there is a relationship between the economic conditions and recoveries.
 - Because GCD data comprises bank initiated not traded loans, the data set differs from most other studies. Hence the outcome can be compared to, but should not be expected to be the same as, studies which focus on publicly recorded bond defaults, single country data or liquidation only data.
- The long term average LGD levels in this report can be compared to regulatory minima and standardised levels, allowing an industry wide discussion of prudent forward looking LGDs vs historical evidence. Note that the LGDs in this report are cash flow discounted observations of historical outcomes, not forward looking estimates.
- Member banks receive the detailed database and can therefore confirm these results and test them on customised sub-sets of the data.

INTRODUCTION

Global Credit Data – established in 2004 – manages the collection of historical LGD, EAD and default observations from over 50 member banks. GCD’s LGD/EAD data set is one of the world’s largest sources of information on all aspects of LGD modelling for wholesale lending. The database comprises over 185,000 defaulted loans to more than 100,000 borrowers covering 11 Basel asset classes.

GCD data is detailed enough to develop or enhance internal LGD models or for validation, calibration or benchmarking. These models can be used to support the Advanced Internal Ratings-Based approach (AIRB), to fulfil the credit provisioning standards IFRS9 or CECL, for stress testing and also for economic capital and pricing.

COMPOSITION OF THE DATABASE AND REFERENCE DATA SET CREATION

GCD recognises that there are different aggregation levels used by its members and therefore two reference data sets (RDS) are used in this study:

- Large Corporates (LC) aggregated on obligor level where loans for each borrower are aggregated
- Large Corporates aggregated on obligation level where each loan or facility is treated separately

Results are shown on both levels in the tables, even though the graphs are on obligor level. Obligation level results are so similar that they are not repeated in the graphs. Individual facility (loan) outcomes do vary greatly for each borrower, which may depend on contract conditions, collateral differences, laws or even bank policies. However, across many borrowers the outcome becomes even, both per year and over time.

Most of the facilities in the GCD datasets are loans of some type, e.g. term loans or overdrafts, however the data also includes significant numbers of contingent facilities, e.g. letters of credit or derivatives as well as some bonds and equity.

**TABLE 1
REFERENCE DATA SET 2019**

	Number of Borrowers	Number of Facilities
Reference Data Set (RDS)	10.737	18.465
RDS Unsecured	4.029	7.917
EAD	196 bn EUR	
Number of Lenders	58	

Only resolved defaults are included in the Reference Data Set (RDS), i.e. cases for which the workout is completed, submitted until November 2018. Cases with year of default from 2000 to 2015 were chosen due to completeness, one more year more than last report. Pre-

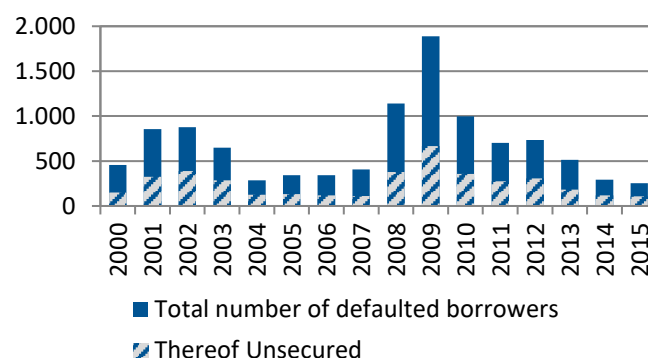
2000 defaults can be biased due to limited and therefore not representative data, while post 2015 defaults contain too high a mix of quick workout (cure) cases. For a detailed description of the RDS creation see the Appendix. The RDS contains 10,737 defaulted borrowers and 18,465 facilities (see Table 1). These defaults are from 58 lenders worldwide (three more lenders than in the previous report).

A significant portion of the data comes from completely unsecured situations (see definition in Appendix) which allows for a more detailed analysis.

The LGD in this report are observations of historical outcomes and not forward looking estimates. They do not reflect specific portfolio alignment or addition of any statistical uncertainty add-ons.

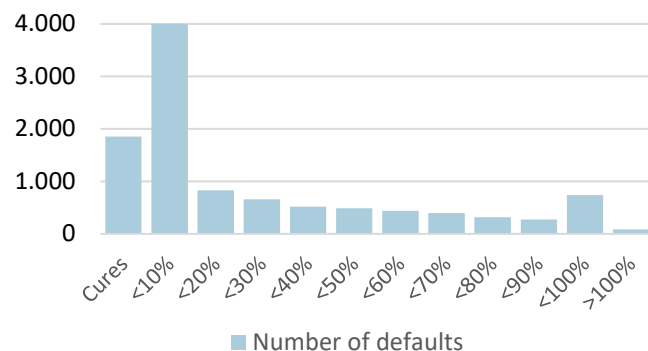
The data covers two clear downturns (see Exhibit 1), 2001-2002 tech stock crisis, and 2008-2009 global financial crisis.

**EXHIBIT 1
BORROWERS BY YEAR OF DEFAULT**



A well-known characteristic of LGD is the bimodal left-skewed distribution (see Exhibit 2) which generates large variations when calculating average LGD. Note in the graph that the cures, which by definition have a nominal LGD of zero, are displayed separately from the LGD bucket <10%.

**EXHIBIT 2
DEFAULTS BY LGD BUCKETS AND CURES**



Banks recover from defaulted loans either most of the outstanding loan amount or almost zero. Receiving a partial repayment of the outstanding amount is less likely to be observed than observing either of these extremes, although it does occur. Indeed, when an average LGD is derived from an RDS the calculated average LGD value usually falls into an LGD bucket which exhibits a lower frequency of loans such as: 20% to 30% or 30% to 40%. The bimodal distribution has implications for measures of spread such as standard deviation.

A simple standard deviation calculation will produce extreme values and larger amounts of data are required to stabilise the central tendency. The variation of the mean is shown here by bootstrap confidence intervals: a simple non-parametric method for constructing confidence intervals.

LGD STABILITY AND RDS COMPARISON

The stability of the 2019 RDS is assessed from two different perspectives. First, the effect of using the same RDS filtering as 2018 is analysed, assessing the overall changes in the data (see Table 2 and Exhibit 3) and the resulting LGD averages and trends (see Exhibit 4). Subsequently, the impact of adding more recent years of default is analysed separately (see Table 4). This analysis of newly arrived data follows the same methodology that banks must use every year on their internal data.

In Table 2 differences between 2018 RDS and 2019 RDS are reported, to allow for a detailed comparison. The 2019 RDS comprises three new lenders, over 1000 more borrowers with more than 1500 loans ([LGD report Large Corporates 2018](#)).

TABLE 2
REFERENCE DATA SETS DIFFERENCES

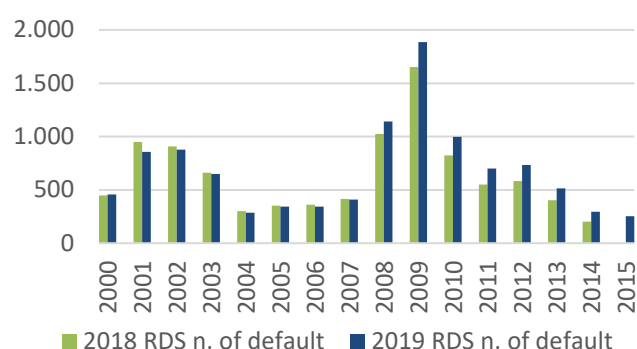
	Number of defaults	Number of facilities
Confirmed from the previous report	9042	15365
One more year of default	252	377
New resolved defaults	1443	2723
Total	10737	18465

There is an overlap of over 80% of the data between the 2018 and 2019 datasets, this indicates the GCD database stability over time. Differences can be categorized as follows:

- inclusion of 2015 year of default (2% of the entire data)
- new resolved defaults. These defaults include defaults submitted by the three new members as well as recently resolved defaults from long-term members. Although the submissions by the three new lenders are intuitively explained, newly resolved

defaults are also submitted by lenders already included in the 2018 RDS. Those defaults are not contained in the previous data set because the RDS filters for resolved defaults where the workout is completed and the ultimate LGD is known. As members continuously work on improving the data quality some defaults could now be included as they pass the data quality filter criteria (see Appendix for more details) and some defaults dropped out of the dataset as they no longer meet the increasingly stringent data quality criteria.

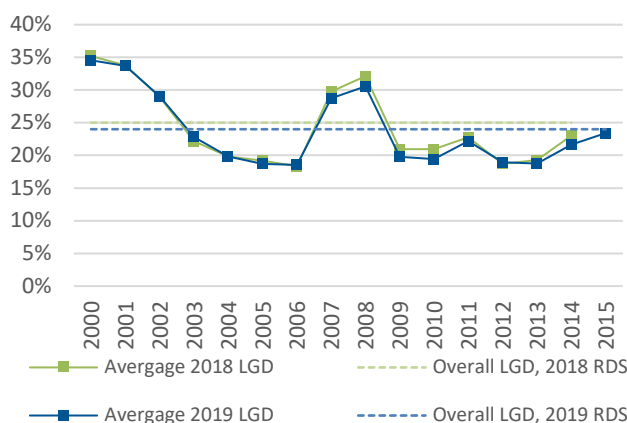
EXHIBIT 3
2019 AND 2018 RDSs



Newly resolved defaults are added to the most recent years, and few defaults are dropped out from the 2001-2002 years. This creates a dilution of the downturn effect with a consequent lower LGD of in absolute terms 1%.

Comparing LGD averages with the previous report there is an overall difference of 1% (22% vs 23% LGDs for Secured obligors; 27% vs 28% LGDs for Unsecured obligors). Exhibit 6 reports the 2019 LGDs for Secured and Unsecured defaults at obligor level, including the bootstrapping confidence interval. The difference of 1% is well within the range of the confidence interval and therefore can be interpreted as not significant.

EXHIBIT 4
YEARLY AVERAGE 2019 LGD VS 2018 LGD



**TABLE 3
SENIORITY AND COLLATERAL**

	Obligor level				Obligation level			
	Number of defaults	LGD	Time to Resolution [years]	Time to Recovery [years]	Number of facilities	LGD	Time to Resolution [years]	Time to Recovery [years]
Total Secured	6,708	22%	2.0	1,2	10,548	22%	1.9	1.2
Total Unsecured	4,029	27%	2.0	1,3	7,917	25%	2.0	1.2
thereof Senior	3,636	26%	2.0	1,3	7,317	24%	2.0	1.2
thereof Subordinated	109	38%	2.0	1,3	227	36%	2.2	1.2
thereof Other/Unknown*	284	35%	2.3	1,5	373	37%	2.4	1.6
Total	10,737	24%	2.0	1.2	18,465	23%	2.0	1.2

* Borrowers are not always borrowing uniquely senior or subordinated. Occasionally a bank will provide facilities of differing seniority to the same borrower. We also include the small number of bond and equity defaults as well as unknowns here.

Exhibit 4 visualizes the impact of adding newly resolved default to the sample on the yearly average LGDs. As expected the differences are all the more negligible the more the data is in the past. Due to the typical resolution time patterns (see Exhibit 8) more data is added to more recent default years. Together with the quick workout cases, this motivates the exclusion of the most recent default years from the RDS. In summary the distribution over time do overlap, suggesting a not significant difference between them.

**TABLE 4
LGD STUDY PERIOD**

years of default	Obligor Level	Obligation Level
	LGD	LGD
2000-2015	24%	23%
2000-2014	24%	23%
2000-2013	24%	23%
2000-2012	24%	24%
2000-2011	25%	24%
2000-2010	25%	24%

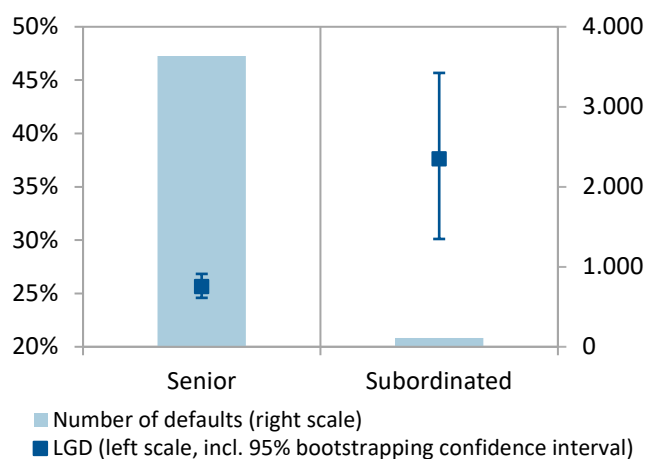
The analysis in Table 4 shows that adding one more year of default does not lead to drastic movement in the average LGD for the years from 2010 up to 2015. Given the benign macroeconomic scenario in those years, this stability in LGD is expected.

Putting all of these analyses together, the GCD database shows consistency and reliability over time, important features for detailed default and loss modelling purposes. This stability also holds for the discriminatory factors of seniority, secured collateral and region. It also holds for the yearly averages, which are especially important in time series analyses used in IFRS9/CECL or stress testing/CCAR exercises.

SENIOR UNSECURED LGD SIGNIFICANTLY LOWER THAN SUBORDINATED UNSECURED

LGD depends on seniority and collateral, LGD outcome is usually lower for collateralised defaults. GCD data confirm that secured LGDs are lower than unsecured (i.e. 22% vs. 27% on obligor level). Unsecured defaults are a relatively homogeneous data set that should isolate the impact of seniority. The data endorses the strong impact of the rank. As in 2018 senior unsecured defaults have a significantly lower LGD than subordinated unsecured defaults (see Exhibit 5).

**EXHIBIT 5
SENIOR AND SUBORDINATED UNSECURED LGD**



When analysing secured defaults, the characteristics of the collateral are expected to strongly influence the LGD outcome. Examples include type of the collateral, the collateral value and the Loan to Value ratio. Therefore, any analysis on secured LGDs needs a deeper view including the above-mentioned items. GCD members can choose from a large selection of extra fields of collateral detail, for further analysis.

TABLE 5
LGD BY REGION

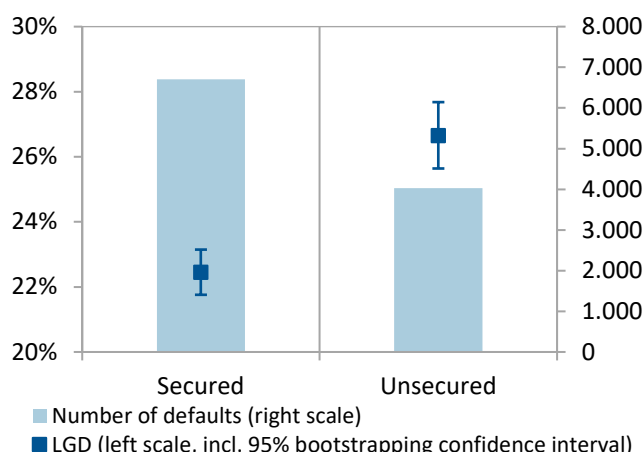
	Obligor level				Obligation level			
	Number of defaults	LGD	Time to Resolution [years]	Time to Recovery [years]	Number of facilities	LGD	Time to Resolution [years]	Time to Recovery [years]
Africa & Middle East	276	20%	2.1	1.5	406	18%	2.1	1.4
Asia & Oceania	1,000	31%	1.8	1.2	1,776	28%	1.8	1.1
Europe	3,831	21%	2.1	1.2	7,459	20%	2.1	1.2
North America	4,444	23%	2.0	1.3	6,847	24%	2.0	1.2
Latin America	1,107	30%	2.1	1.3	1,832	29%	1.9	1.3
Unknown	79	44%	3.4	1.8	145	44%	3.7	1.7
Total	10,737	24%	2.0	1.2	18,465	23%	2.1	1.2

SECURED LGD LOWER THAN UNSECURED LGD

On a single driver analysis, consistently over time, GCD data shows that collateralised loans and obligors produce significantly lower LGD than unsecured loans and obligors. This outcome supports common bank lending policies which assume that the taking of collateral will improve the bank's position versus unsecured creditors.

As shown in Exhibit 6, this effect is limited to 5%, as in the last year's report. One of the reasons may be these same policies, which often require that less capitalised companies, with more volatile assets are more likely to be required to give collateral to the lending bank. Thus, the lending bank compensates for expected weak recoveries and higher LGDs by taking collateral to improve this.

EXHIBIT 6
SECURED AND UNSECURED LGD



TIME TO RECOVERY AND RESOLUTION

Time to resolution is on average around 2 years. Generally, a default can resolve because of three reasons. First, the borrower pays back all the debt. Second, the borrower returns to a non-defaulted status. Third, the bank decides to stop the recovery efforts and writes off

the outstanding debt (or sells it). Only the first option depends entirely on the borrower while the other two are influenced by choices made by the bank involved. Continued forbearance is also under the control of the lending bank.

We therefore apply a different measure of the time in default that is more objective as it only depends on the timing of the cash flow. The Time to Recovery (TTRec) is the average period between default and cash flow payment weighted by the amount of the payment. TTRec is by definition shorter than or equal to the time to resolution. The outcome is remarkably similar for TTRec across differing collateral and seniority states.

As the TTRec represents the middle point of the cash flows, it is a good measure for understanding the effect of discount rates on the LGD. The cash flows are discounted on average for approximately 1.2 years and not the full 2 years of the average Time to Resolution.

REGIONAL VARIATIONS

The country or region of the borrower is expected to be a driver of LGD, as lending practices, insolvency laws and regional economic differences should affect recoveries. The GCD data set offers country information on several levels: country of residence; country of jurisdiction; collateral country of jurisdiction. The impact is best analysed on country level, but granularity must be weighed against availability of a significant amount of data points. Reflecting the global membership base of GCD, there are almost 140 countries reported in the data.

This study shows the LGD by region based on the country of residence of the defaulted borrower.

The data set in this report comes from Large Corporate borrowers, defined by their sales or assets being above €50m. Many of these companies have multi country

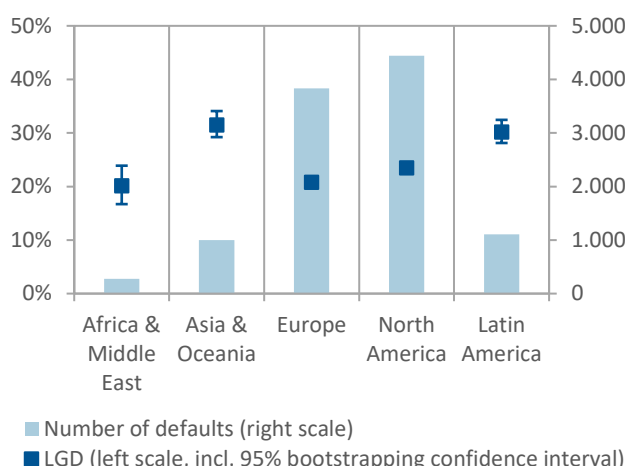
operations and participate in cross border trade, which could act to reduce the regional variation.

EUROPEAN AND NORTH AMERICAN DATA SHOW SIMILAR RESULTS

GCD data has its strongest database in Europe and North America, which register similar figures. Table 5 appears to show Africa and the Middle East with slightly lower LGD, however this is based on a much smaller data set compared to the other regions (see Exhibit 7) and is not yet significant.

Asia and Oceania appear to have a higher LGD. This data is comprised by a large number of different countries – most notably Australia, South Korea and Hong Kong. The country-level LGDs for Australia, Japan, South Korea and New Zealand are more or less in line with the European and North American data, but other countries show higher LGDs – pushing up the region’s average.

**EXHIBIT 7
LGD BY REGION**



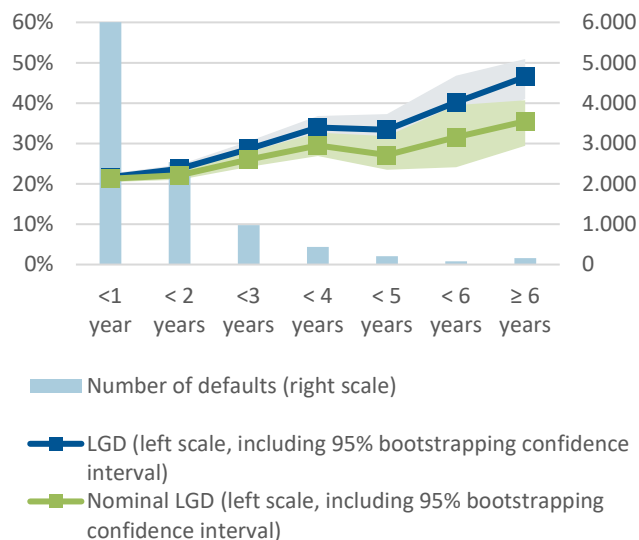
HIGHER LGD FOR LONGER TIME TO RECOVERY

Consistently with last year results, Exhibit 8 shows a clear trend on how workout period correlates with the LGD outcome. Average LGD levels based on TTRec buckets are displayed. The longer the TTRec the higher the LGD. Because this effect can be related to higher discounting effects in addition to the usual LGD which is discounted, nominal LGDs were added in the picture (green line).

The nominal LGD is lower than the discounted numbers but still rises steadily. It may be that workouts which take more time due to their complexity, legal disputes or other factors just happen to have lower recoveries and higher LGD. Alternatively, banks may proceed more quickly in recovering cases where the prospects are strongest. TTRec is shown as it best relates to the cash flow timing. The distribution of the time to recovery buckets

complements the information on the averages displayed in Table 5 and 6.

**EXHIBIT 8
LGD BY TIME TO RECOVERY**



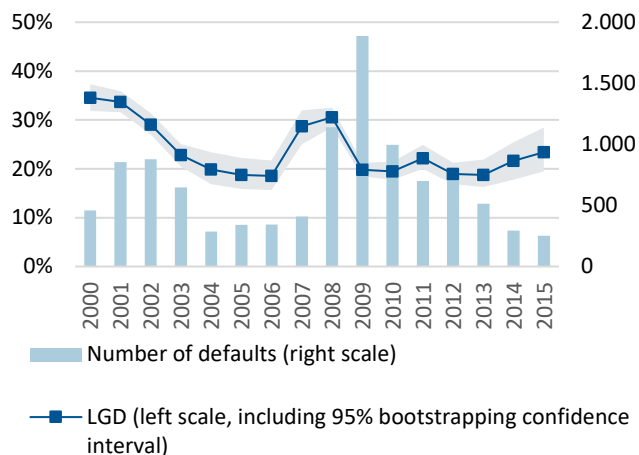
For most cases the main cash flows occur in the first year after default. The number of defaults per bucket decrease steadily. Nevertheless, there are a number of data points with an over 6-year average recovery period which are grouped into one bucket.

MACROECONOMIC EFFECTS IN HISTORICAL LARGE CORPORATE DATA

The observation of downturn effects in historical LGD data is typically complicated by short time series, few data points and the multitude of input parameters for LGD estimates. Requirements like IFRS9/CECL or stress testing/CCAR create the need for more detailed default and loss modelling, especially in respect of term structure and macroeconomic dependency. The long timespan of the GCD database and the detailed cash flow data allow for dedicated LGD time series analysis. To assess the variance over time the LGDs are plotted by year of default in Exhibit 9, with the corresponding numbers displayed in Table 6. The shape of the curve shows variance over time with higher LGDs in the early 2000s as well as in 2007/2008.

Results are consistent and in line with 2018 report, highlighting a strong stability over time for GCD database. GCD has extensively analysed downturn effects on LGD especially including the distribution of cash flows over time. The recovery cash flows are dispersed over varying periods of time. On average the workout period lasts 2 years but recoveries can be collected over a much longer period (see Exhibit 8) which is even longer if excluding cures.

**EXHIBIT 9:
LGD BY YEAR OF DEFAULT**



Looking at the timing of the underlying cash flows, the evolution of loss given default values over time can be analysed with respect to their co-movement with macroeconomic indicators. The results are published in GCD's [Downturn LGD Study 2017](#).

Table 6 displays the volumes and LGD averages by year of default, aggregated at both obligor and obligation levels. Not surprisingly, the difference between these levels is small.

**TABLE 6
LGD BY YEAR OF DEFAULT**

Year of default	Obligor level		Obligation level	
	Number of defaults	LGD	Number of facilities	LGD
2000	458	35%	743	35%
2001	855	34%	1,421	32%
2002	878	29%	1,366	27%
2003	648	23%	1,034	22%
2004	286	20%	480	17%
2005	342	19%	560	19%
2006	344	19%	526	19%
2007	408	29%	751	30%
2008	1,141	31%	1,965	30%
2009	1,887	20%	3,457	19%
2010	997	19%	1,767	19%
2011	701	22%	1,270	22%
2012	733	19%	1,384	21%
2013	513	19%	851	19%
2014	294	22%	513	21%
2015	252	23%	377	23%
Total	10,737	24%	18,465	23%

CONCLUSIONS

In summary, conclusions can be drawn from the analyses presented here regarding the following questions:

- **Is GCD Data consistent over time?** The stability analyses presented here show consistency over time for GCD Large Corporate data. The overall LGD averages as well as the yearly averages show only small and not significant differences of in absolute terms 1%. Trends and results through all the report, including the other metrics (Time to Resolution and Time to Recovery) are also very similar to last year's results.
- **Does LGD differ if it is calculated for each facility or at overall borrower level?** This report confirms the average LGD level over time differs by only 1% for this factor.
- **What drives LGD?** Seniority and collateral are confirmed as LGD drivers. Secured LGD is lower than unsecured LGD (22% vs 27% on obligor level, 22% vs 25% on obligation level). For unsecured LGD, seniority is confirmed as a driver at obligor level (26% senior vs 38% subordinated) and at obligation level (24% vs 36%).
- **Are regional variations significant for LGD?** After aggregating country level data to regions, North America and Europe appear to have similar levels of LGD.

The insights gained by the high-level results presented here confirm the benefit of a detailed and granular collection of post default cash flow data. It is crucial for banks using a data-driven credit risk estimation method to understand and quantify loss given default.

NOTE ON METHODOLOGY

This report is based on a certain LGD methodology. Definitions used, calculations made and data filters applied are laid out in detail in the **Appendix** to this report. Evidence for the consistency and veracity of the GCD data is presented. GCD members have access to the raw but anonymised data which enables them to produce customised representative data sets and calculate averages suitable for their own portfolio comparisons.

OUTLOOK

For this study, large corporate data was used. The same analytics can be performed on the other asset classes in the GCD data.

GCD provides to its member banks extensive peer comparison reports, where the portfolio composition and LGD outcomes are contrasted with the defaults of the

other members. This enables them to answer the question of where does each bank stands compared to its peers.

NOTE ON TERMS USED

LGD refers to Loss Given Default rate which is calculated as $(1 - \text{recovery rate})$. The recovery rate is the net of all cash flows including external costs (discounting the cash flows at a risk-free rate of 3 months EURIBOR) divided by the amount outstanding at default. The LGD calculation is made using a cap of 150% and floor of 0% using GCD's LGD2 method (see Appendix) where the EAD is increased by the amount of any post default advances. Variations could include using a different discount rate based on a combination of the risk free rate and a risk premium for systematic risk at the time of default.

Nominal LGD is calculated in the same way but using nominal cash flows, i.e. not discounted.

Time to Resolution (TTR) is calculated as the period between the date of default and the date of resolution (i.e. repayment, write-off, return to performing, etc).

Time to Recovery (TTRec) is defined as the cashflow weighted average period between default and cashflow.

A more detailed definition is given in the Appendix.