Global Credit Data

*by banks for banks*

# Isn't there really enough data to produce good LGD and EAD models ?

Dr. Massimo Cutaia

# Contents


- About Global Credit Data and why we are interested
- IRB constraints proposed by Basle committee
- How much data is required?
- How much data is available?
- Examples of use of available GCD Data
- Conclusion


Disclaimer: The views expressed in this presentation are those of the presenter and do not necessarily represent the views of Global Credit Data.

# About Global Credit Data


Our Mission is to help banks understand and model their credit risks.

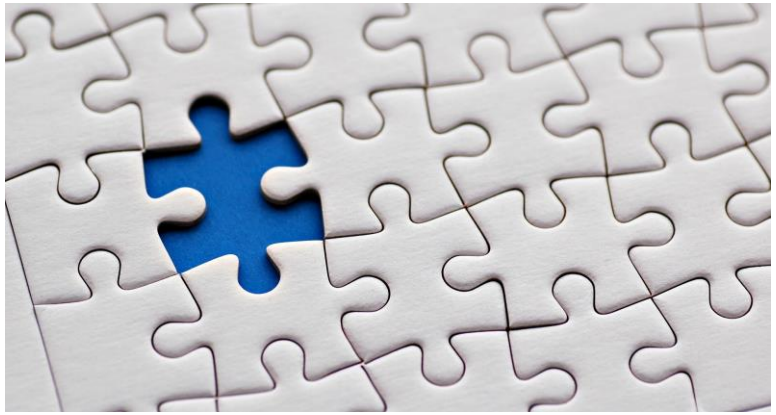
We are a non-profit association owned by (currently) 52 banks.

Our Activities include pool credit loss data, especially for low default portfolios; provide benchmarks; facilitate knowledge exchange and foster research.

| | | | | |
|---|---|---|---|---|
| World’s largest loss databases | Banks in the driver’s seat | Confidentiality | A venue for best practice sharing | Model Benchmarking |

Your presenter: Dr. Massimo Cutaia

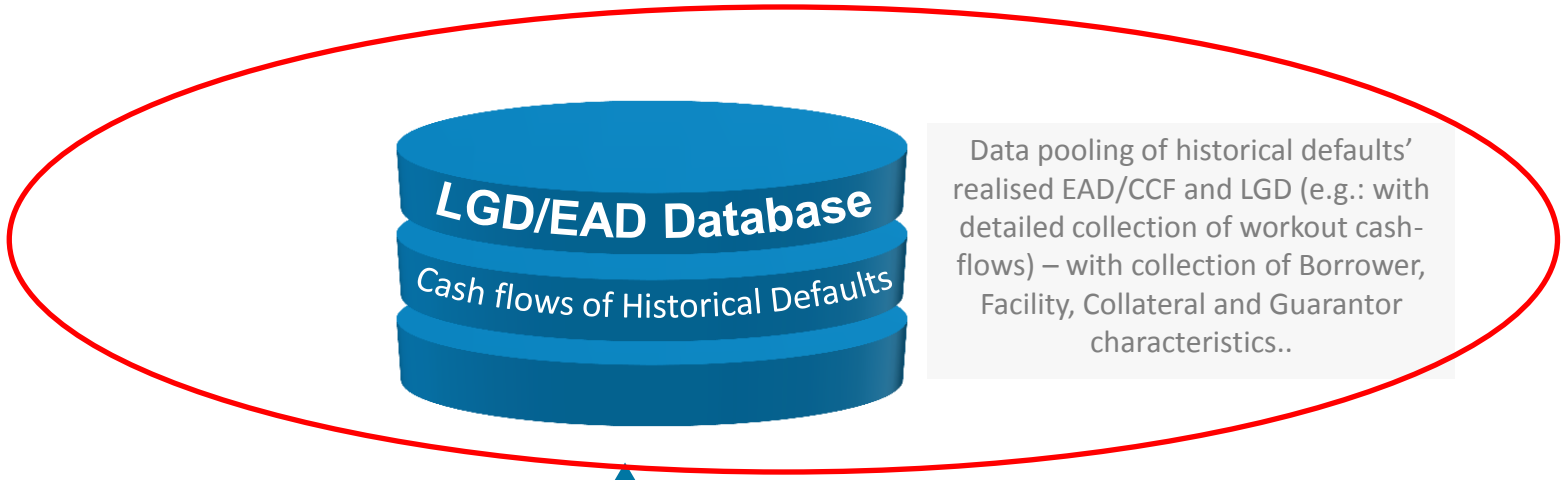
- Global Head of A-IRB modelling at Credit Suisse AG
- Since 10+ years in credit risk modelling
- Member of the board at Global Credit Data & Treasurer





Global Credit Data

# Overview of GCD solutions




Data pooling of historical defaults' realised EAD/CCF and LGD (e.g.: with detailed collection of workout cash-flows) – with collection of Borrower, Facility, Collateral and Guarantor characteristics..

GCD Solutions


Benchmarking of current predicted PD, EAD/CCF and LGD for named counterparties and specific banking book clusters.


Data pooling of historical realised rating migrations and predicted PDs by: asset classes (Large Corp., SMEs, Banks & FIs..), Regions and Sectors..

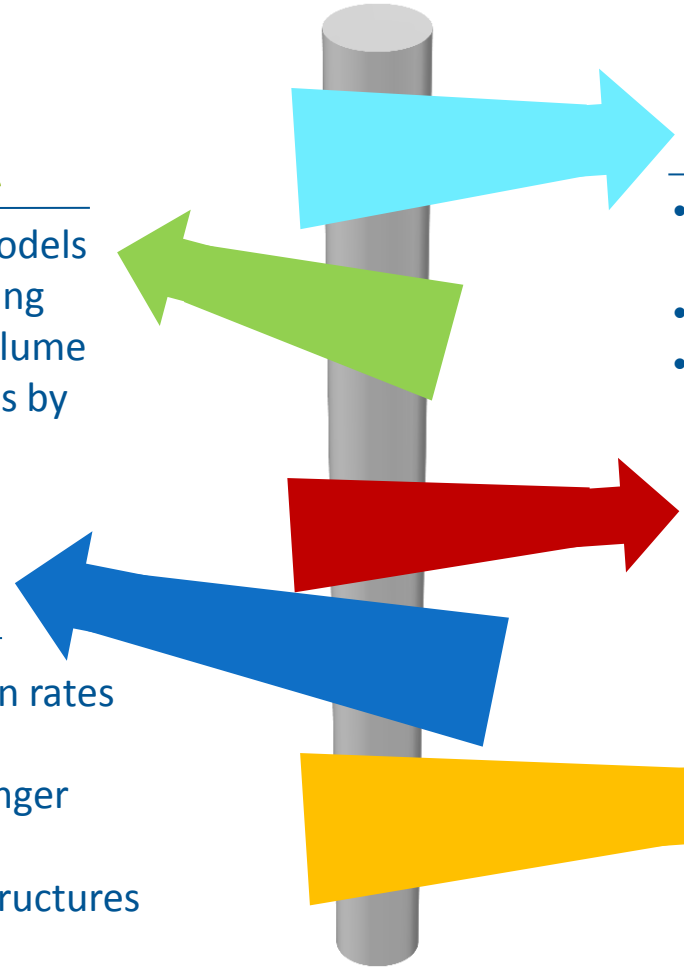




Global Credit Data

# Models are not only for RWA, in other areas there is no fall-back solution

Our **data** is used as a modelling source or for benchmarking/validation of non retail portfolios



## Stresstesting / CCAR

- Build more detailed stress testing models
- Decrease uncertainty add-ons by using longer time series and more data volume
- Assist macroeconomic factor analysis by using our rating transitions

## Regulatory capital

- Prove the correct PD and LGD levels for Low Default Portfolios (e.g. banks, shipping)
- Reduce uncertainty add-ons for lack of data
- Correctly calibrate downturn LGD from long time series

## IFRS 9/ CECL

- Use our multi year default and migration rates to build bucket models
- Calibrate lifetime expected loss from longer time series
- Extract risk drivers for LGD/EAD term structures

## Economic capital

- Use real bank default data to better calibrate loss distribution models (e.g. correlations)

## Pricing

- Peer benchmark the risk estimates underlying your pricing models

# Contents


- About Global Credit Data and why we are interested
- IRB constraints proposed by Basle committee
- How much data is required?
- How much data is available?
- Examples of use of available GCD Data
- Conclusion




Global Credit Data

# “Not all credit risk exposures are capable of being modelled”*

According to BCBS analysis the following portfolios should not be allowed to be modelled for capital purposes any more due to “lack of appropriate data for risk parameter estimation” and “obtaining reliable estimates of LGDs are even more challenging”*

- Banks and other financial institutions
- Large corporates > 50bn completely and for corporates > 200mn only F-IRB shall be allowed
- Specialized lending at best the slotting approach shall be allowed

One criteria for other portfolios to be allowed to be modelled going forward is data availability described as

“The quantity and quality of relevant data available for the risk or portfolio“*

* BCBS publication d362 – “Reducing variation in credit risk-weighted assets – constraints on the use of internal model approaches”

# Contents


- About Global Credit Data and why we are interested
- IRB constraints proposed by Basle committee
- How much data is required?
- How much data is available?
- Examples of use of available GCD Data
- Conclusion

# How much data is required ?

## Data available:

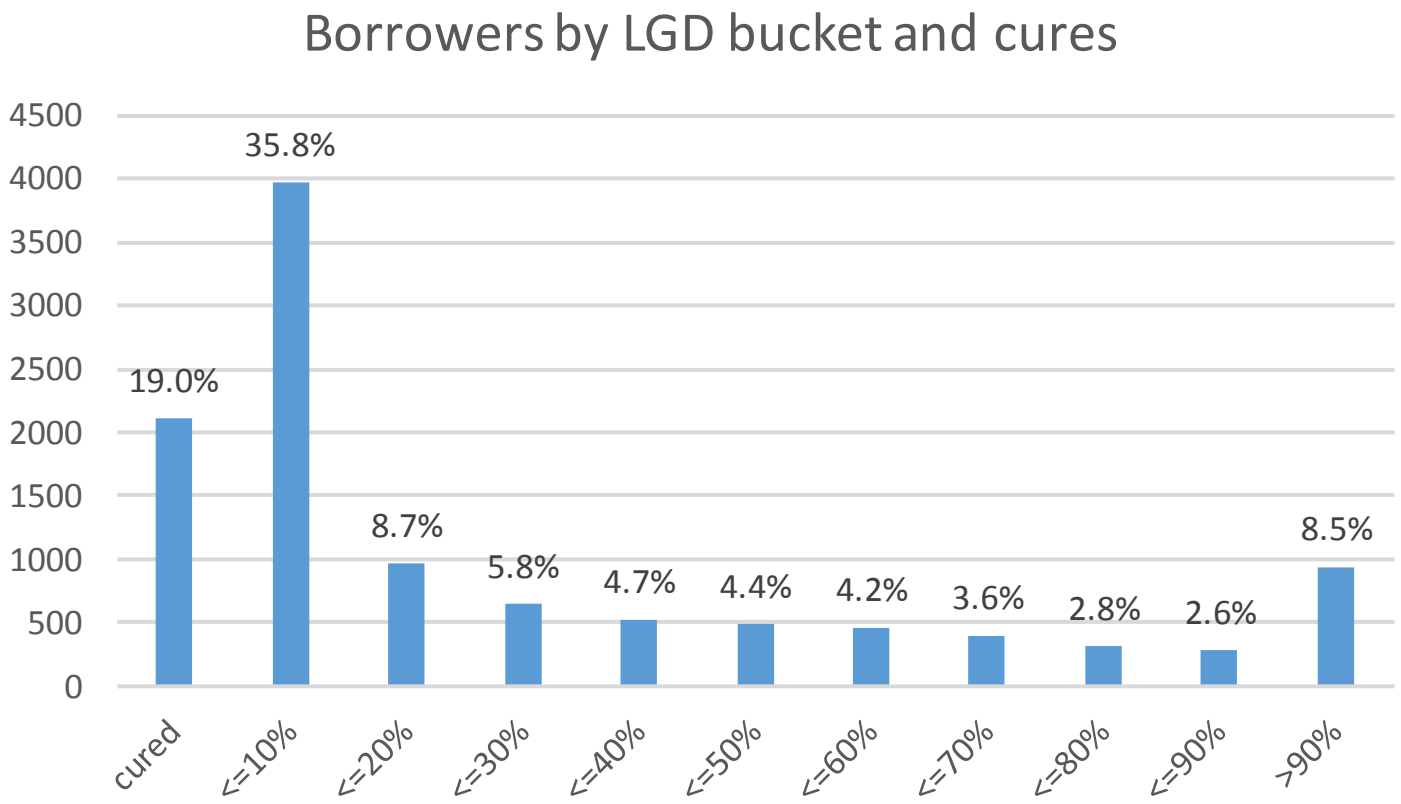
| Hypothetical Bank Portfolio | Retail Mortgage | Large & Mid Corp | Shipping | Sovereign |
|---|---|---|---|---|
| Borrowers | 300,000 | 10,000 | 300 | 100 |
| Defaults per year | 300 | 20 | 3 | 0.25 |
| Loss cases per year | 150 | 15 | 2 | 0.1 |
| Collaterals Sold per year | 50 | 5 | 1 | 0 |

## Sufficient for?

| | | | | |
|---|---|---|---|---|
| PD model build | yes | difficult | unlikely | unlikely |
| PD TTC average level 5 yrs | yes | difficult | low confidence | unlikely |
| EAD undrawn CCF model | yes | difficult | unlikely | unlikely |
| LGD Collateral model | yes | unlikely | unlikely | unlikely |
| LGD general model | yes | unlikely | unlikely | unlikely |
| LGD 7 year average | yes | low confidence | unlikely | unlikely |
| LGD Downturn 20 yrs | yes* | unlikely | unlikely | unlikely |

*if you have 20 years of data

# LGD is bimodal (usually), with many cures




Large Corporates, 2000 to 2014

Note the cure rate, representing 0 nominal LGD in <12 month resolution time and no sale of collateral or call of guarantors

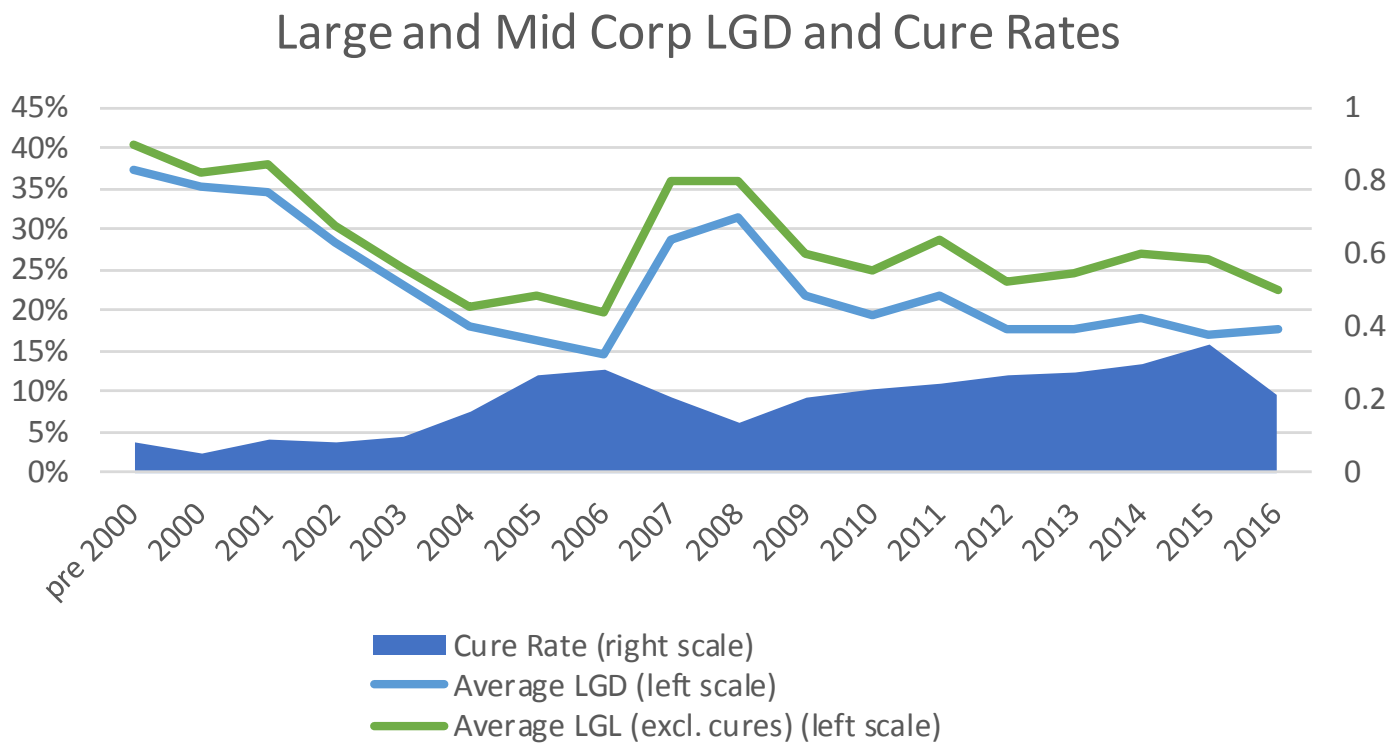
- A problem for LGD is that the observed distribution is bimodal, making the average variable
- Simple standard deviation measures produce extreme outcomes
- Larger amounts of data are required to stabilise the central tendency



Global Credit Data

# Wait for data to mature – avoid the Resolution bias

- LGD is best calculated on closed (resolved) cases
- Cured cases happen in the first year
- Time to resolution for non-cure cases is much longer
- The resolution bias for example leads to unrealistically high cure rates for the most recent years
- Old data may have a lack of cured cases due to delivery issues




Need to cut off recent years or extrapolate the data for yet unresolved cases to reduce the resolution bias needs even more data



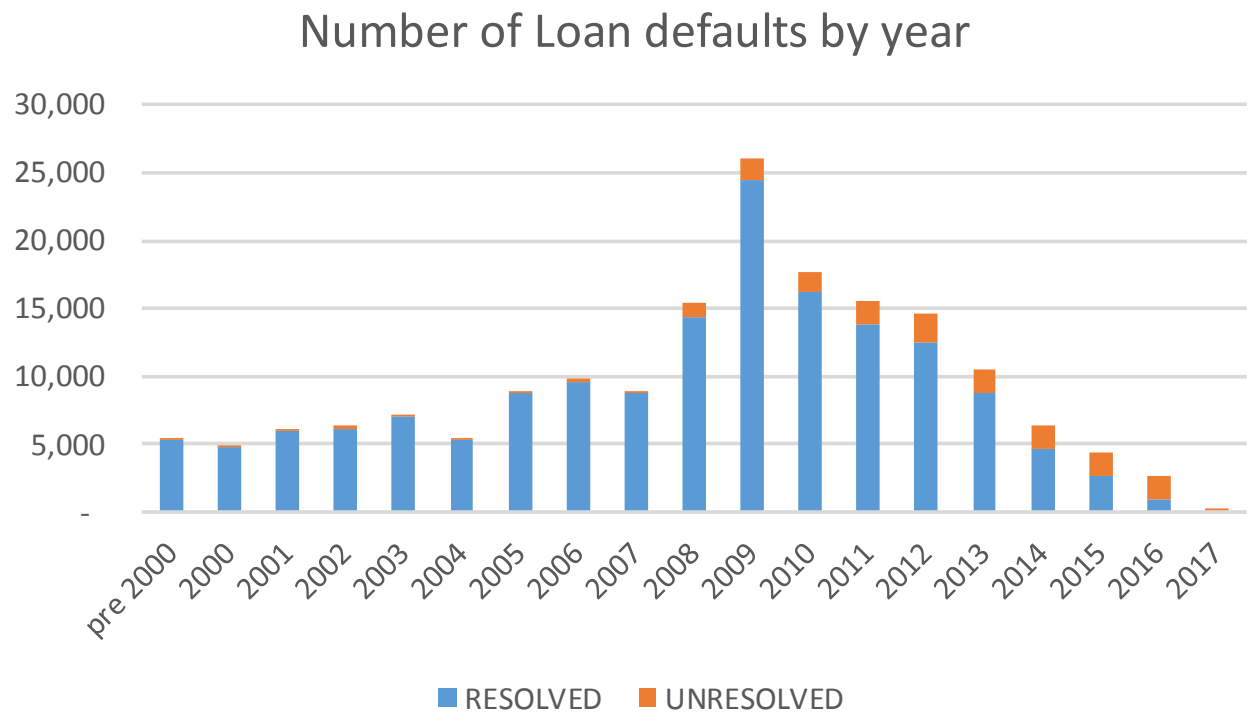
Global Credit Data

# Contents


- About Global Credit Data and why we are interested
- IRB constraints proposed by Basle committee
- How much data is required?
- How much data is available?
- Examples of use of available GCD Data
- Conclusion

# Pooled Defaults in the GCD LGD/EAD Database

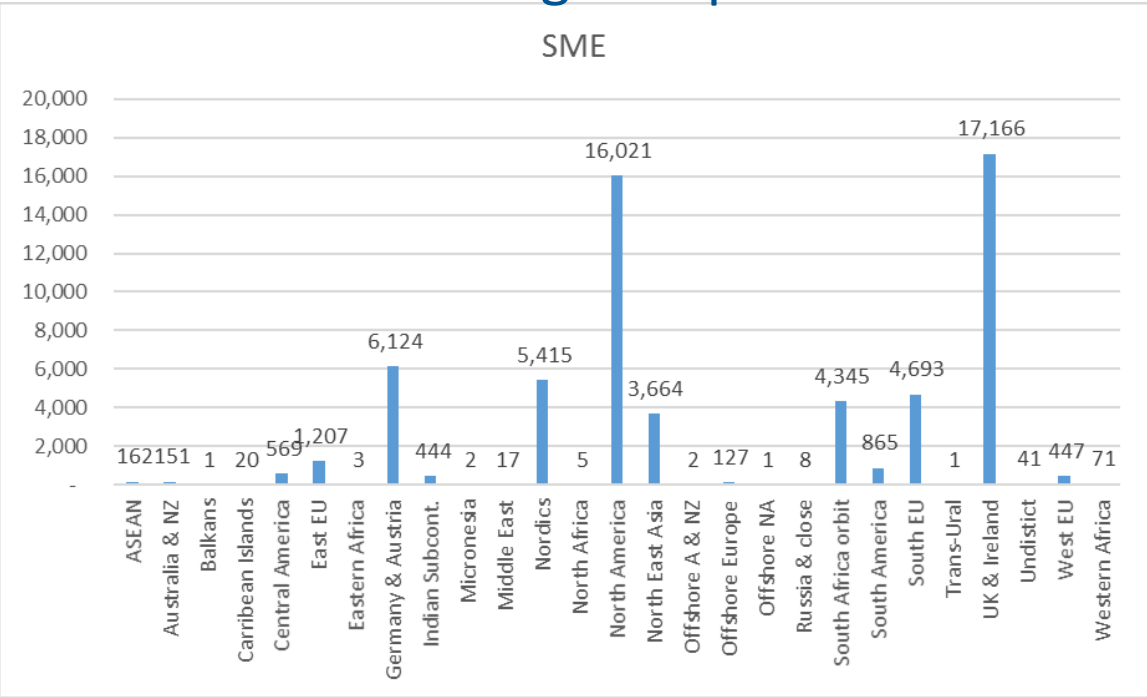
| Facility Asset Class | No. of Defaulted Loans |
|---|---|
| SME | 111,881 |
| Large Corporates | 29,614 |
| Banks and Financial Institutions | 3,623 |
| Ship Finance | 1,208 |
| Aircraft Finance | 801 |
| Real Estate Finance | 20,519 |
| Project Finance | 893 |
| Commodities Finance | 417 |
| Sovereigns & Central Banks | 152 |
| Public Services, Local Authorities | 178 |
| Private Banking | 7,188 |
| **Total** | **176,474** |

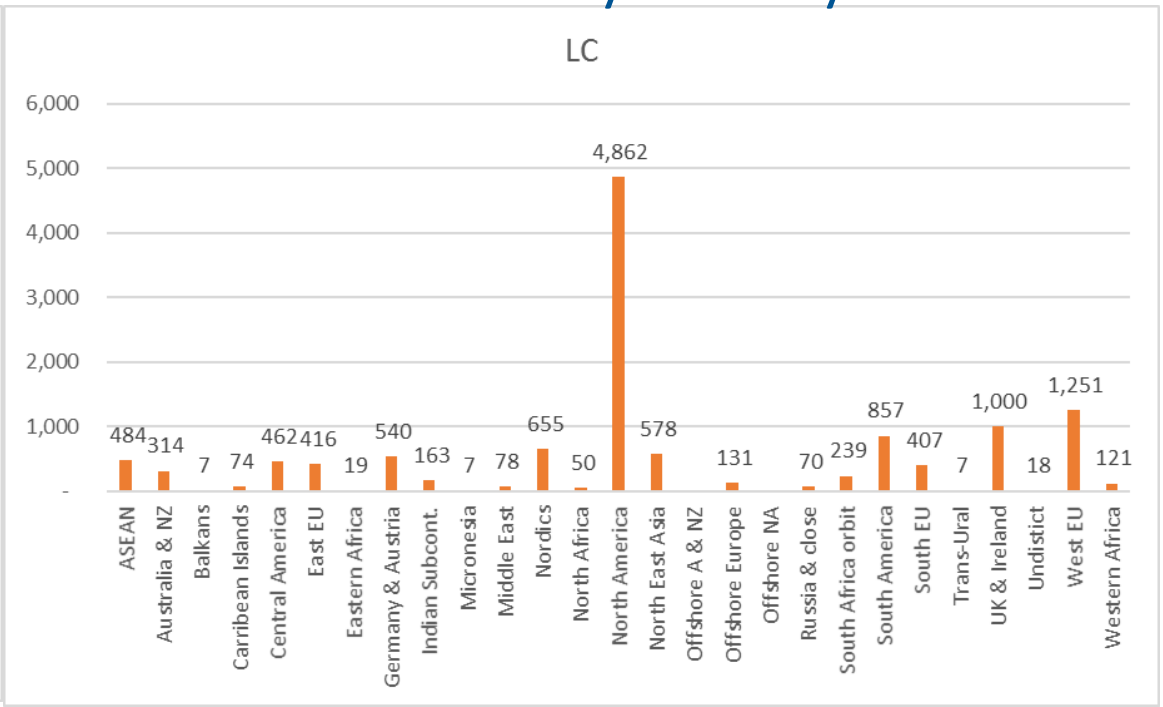



- Over 170,000 loans in total
- Data collected from over 50 banks covering over 20 years
- As expected, SME data is most numerous

# Borrowers in the LGD/EAD Database by region

## Large Corporates and SME Defaulted Borrowers by Country



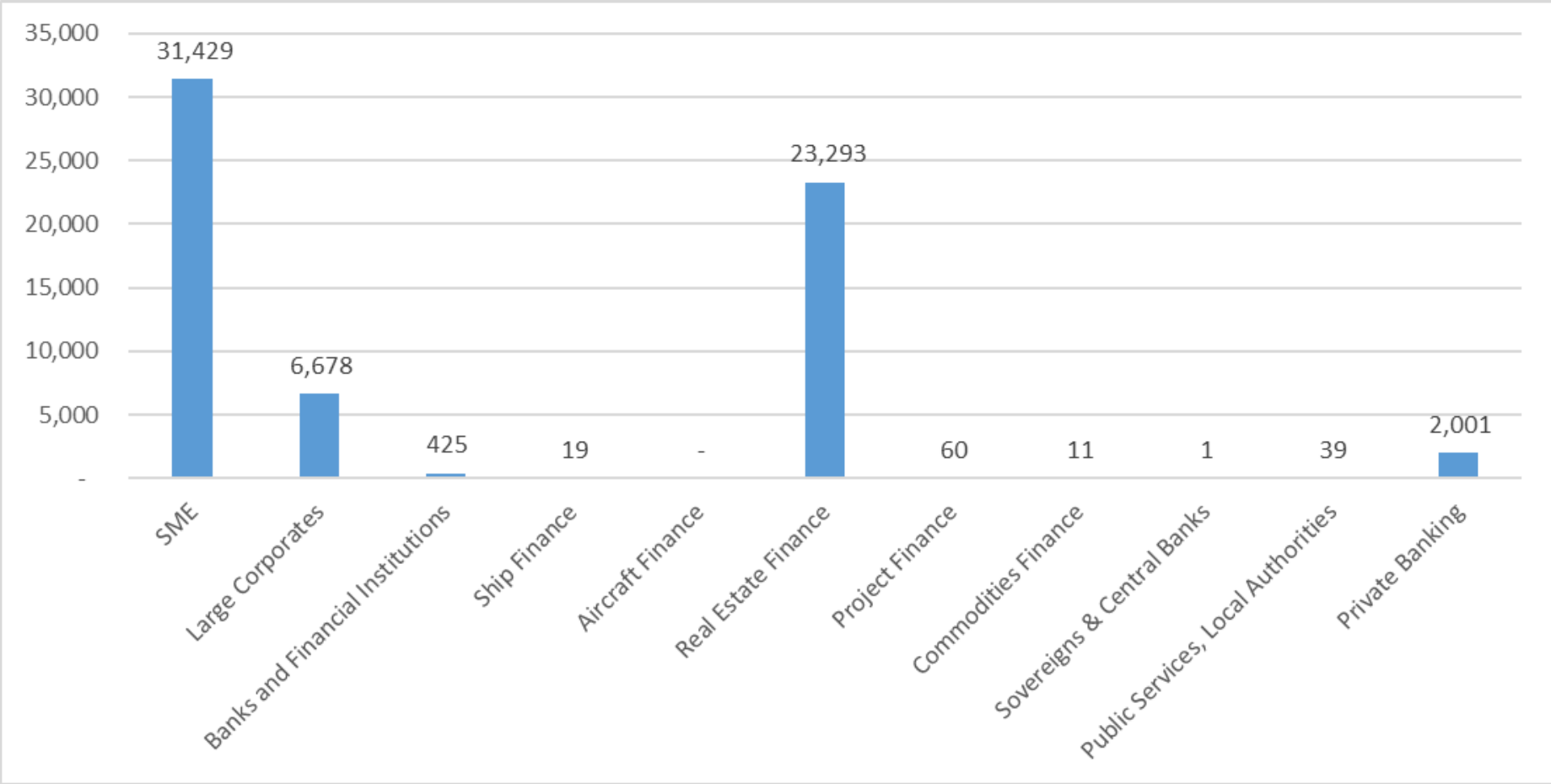

- Information is collected for Country of Residence, Country of Jurisdiction, Country, State and Postcode for Real Estate Collaterals
- Countries are aggregated here, but detailed country information available to member banks



Global Credit Data

# Pieces of Real Estate Collateral in defaulted cases, large amount being specialized lending → IPRE

## Real Estate Collateral by Facility Asset Class





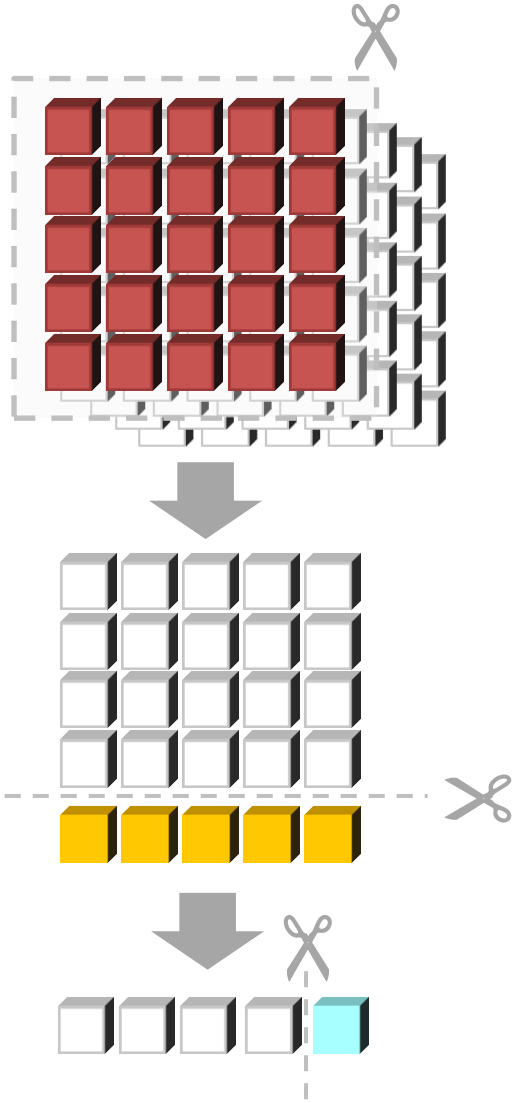

Global Credit Data

# Slicing Pooled Data to a representative set

Pooled data can be used to fill the gaps for low default portfolios, but it must be sliced into representative data sets, fitting the bank's needs:

- Borrower type
- Region or country
- Borrower size (e.g. by turnover)
- Facility type (for CCF)
- Seniority and secured/unsecured
- Collateral type (for secured models)

To achieve a reasonable amount of final data to work with requires a lot of data to start with



# How much data in a pooled data set?

## Data available:

| GCD Data Pools | Retail Mortgage | Large & Mid Corp | Shipping | Sovereign |
|---|---|---|---|---|
| Borrowers | NA | >100,000 | >5,000 | >1,000 |
| Defaults per year | NA | 500 | 40 | 10 |
| Loss cases per year | NA | 300 | 20 | 5 |
| Collaterals Sold per year | NA | 30 | 10 | NA |

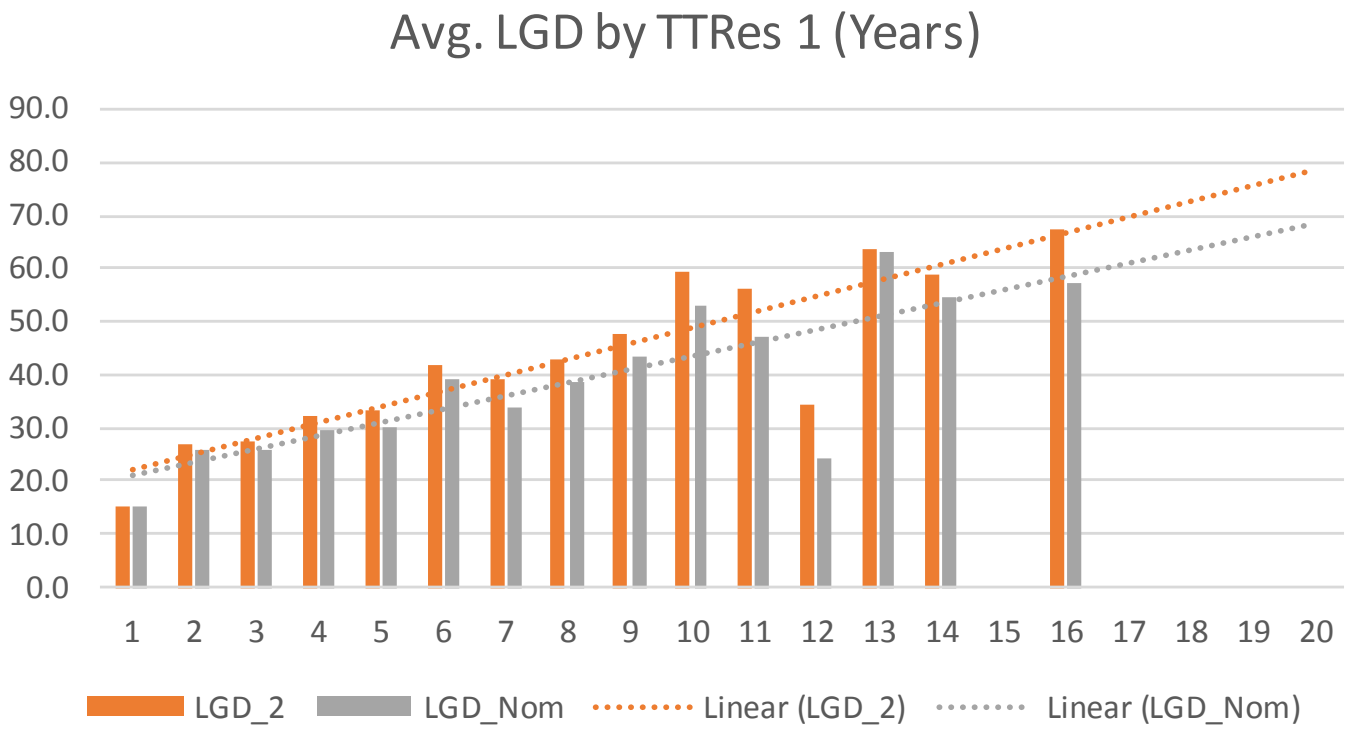
## Sufficient for?

| | | | | |
|---|---|---|---|---|
| PD model build | | yes | simple model | difficult |
| PD TTC average level 5 yrs | | yes | yes | yes |
| EAD undrawn CCF model | | yes | difficult | difficult |
| LGD Collateral model | | yes | yes | difficult |
| LGD general model | | yes | yes | difficult |
| LGD 7 year average | | yes | yes | low confidence |
| LGD Downturn 20 yrs | | yes | yes | low confidence |

# Contents

- About Global Credit Data and why we are interested
- IRB constraints proposed by Basle committee
- How much data is required?
- How much data is available?
- Examples of use of available GCD Data
- Conclusion

# As expected, LGD increases with workout time

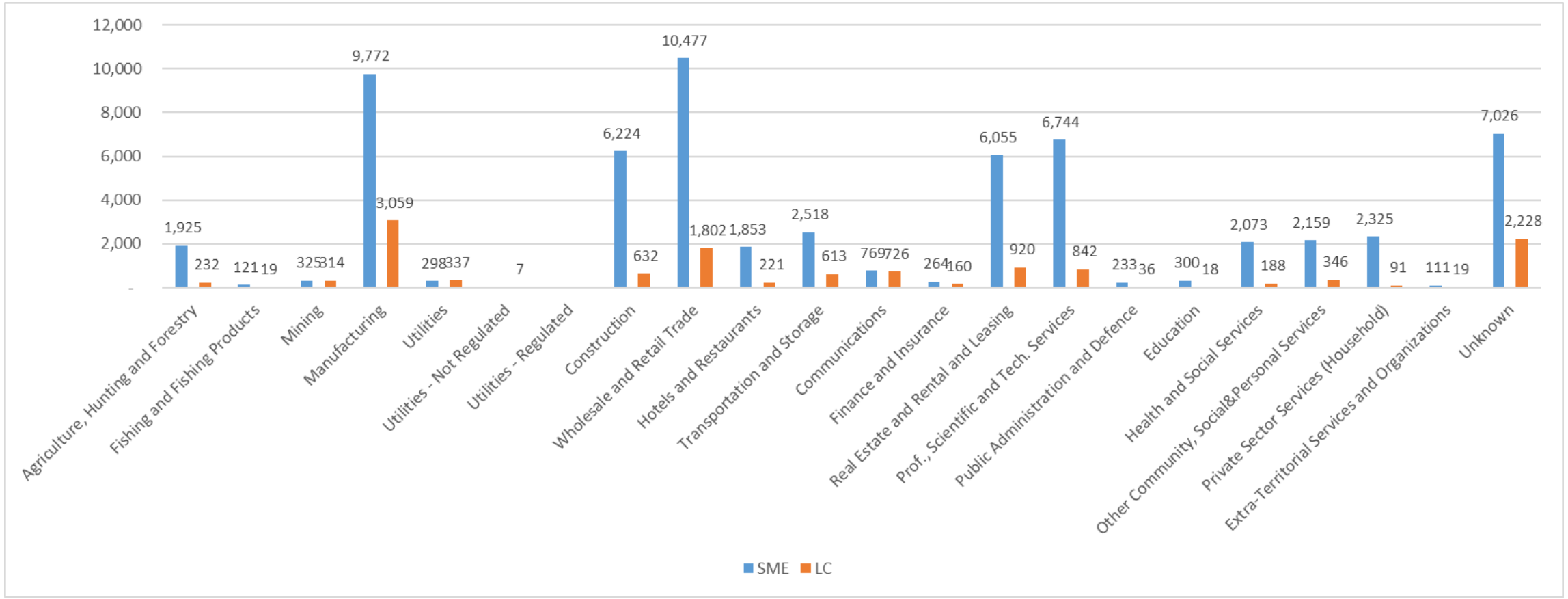



- Large Corporates, 2000 to 2013
- After 1 year of workout all the quick and easy cases are closed
- Remaining cases give lower recovery (higher LGD) in nominal terms and after discounting



Global Credit Data

# Borrowers by industry in the LGD/EAD Database

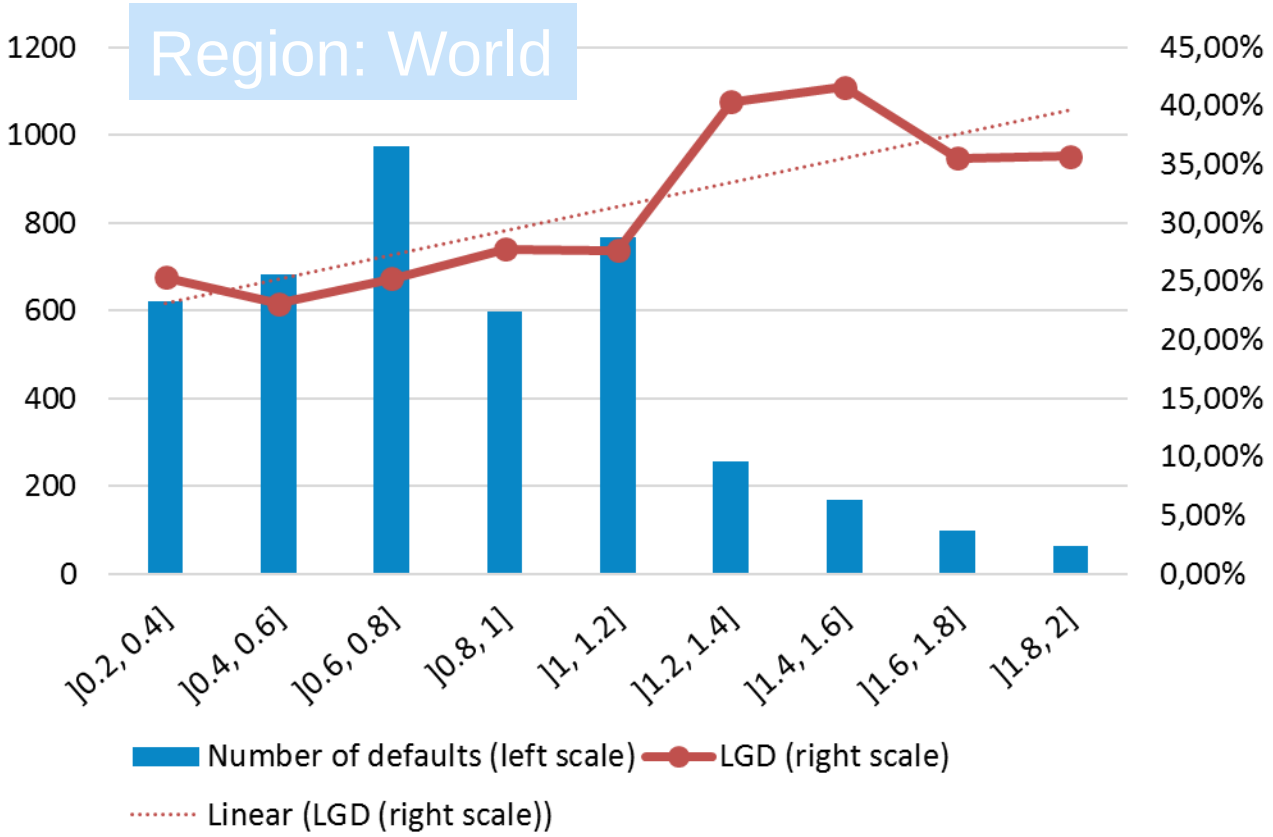



- LGD modellers dream of discrimination by industry
- By merging to 5 or so industry groups, using GCD pooled data, the differences appear



Global Credit Data

# LGD by Loan-to-Value Bucket




*Reference Dataset: H1 2017, Real Estate (Facility Asset Class = Real Estate and/or Collateral Real Estate attached, Real Estate Type = Commercial Real Estate. Resolved. 1 collateral per loan*

- Pooled data shows a positive relationship between LGD and LTV ….but not nearly as strong as expected by regulators
- Of course Real Estate lending risk is much more complex than this and is driven by:
  - Property type
  - Property location
  - Rental contract and tenant strength
  - Cash flow buffers, etc.

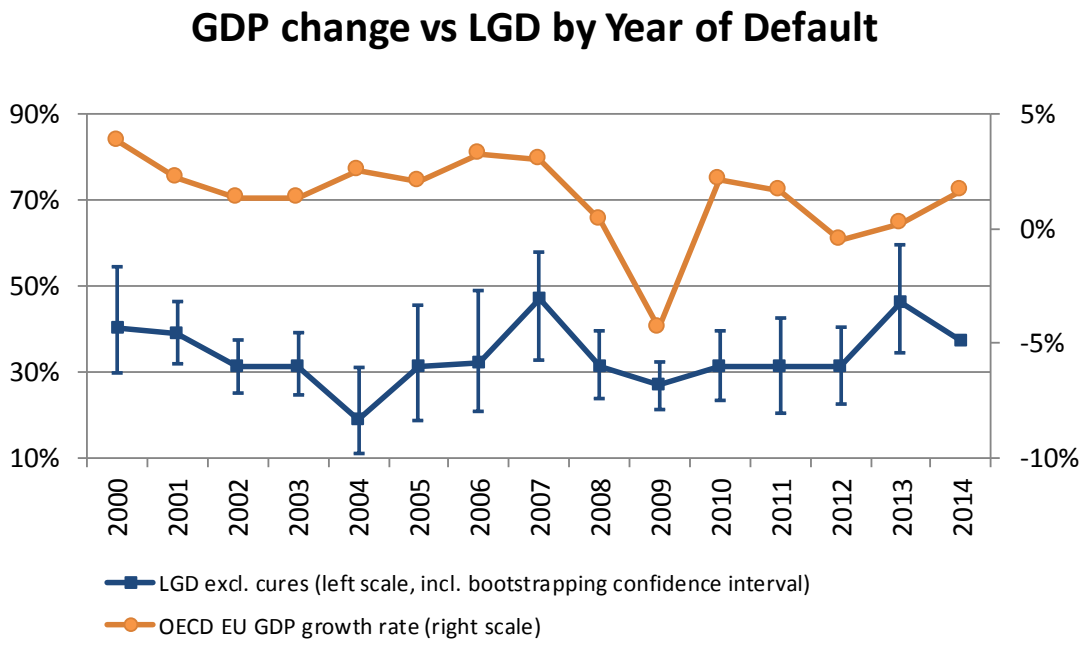
**Note on terms used**

**Loan-to-Value** refers to the ratio of the outstanding amount of a loan to the value of the collateral between one year prior to default and the default date.

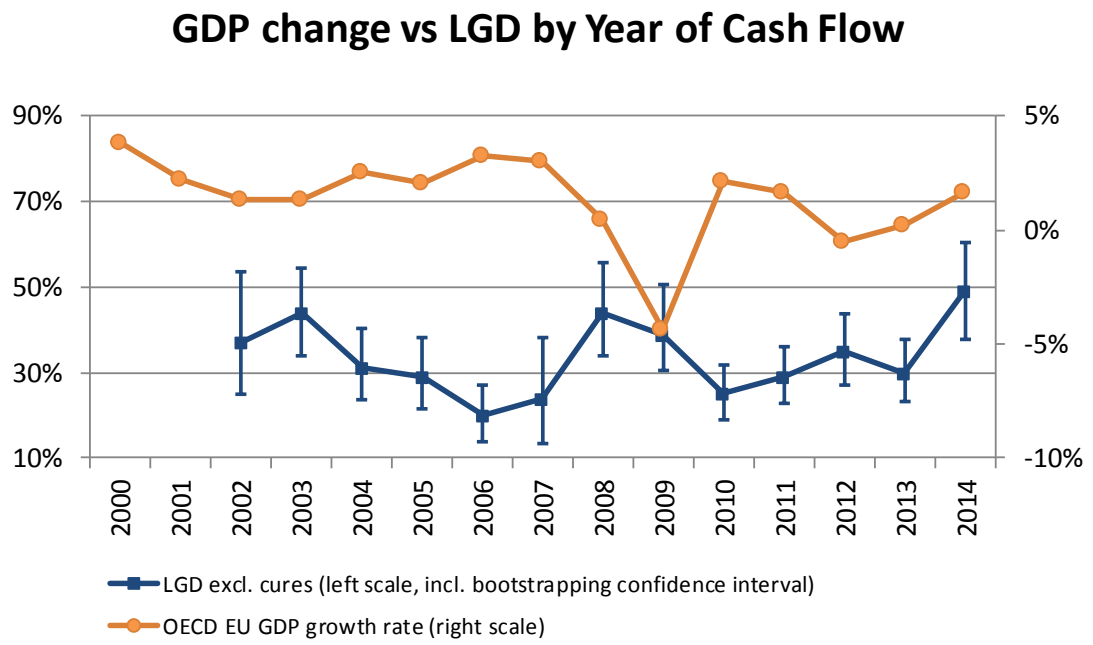


Global Credit Data

# Downturn analysis: Yr of Default vs Yr of Cash Flow




*Asset Class: Large Corporate unsecured, Europe*



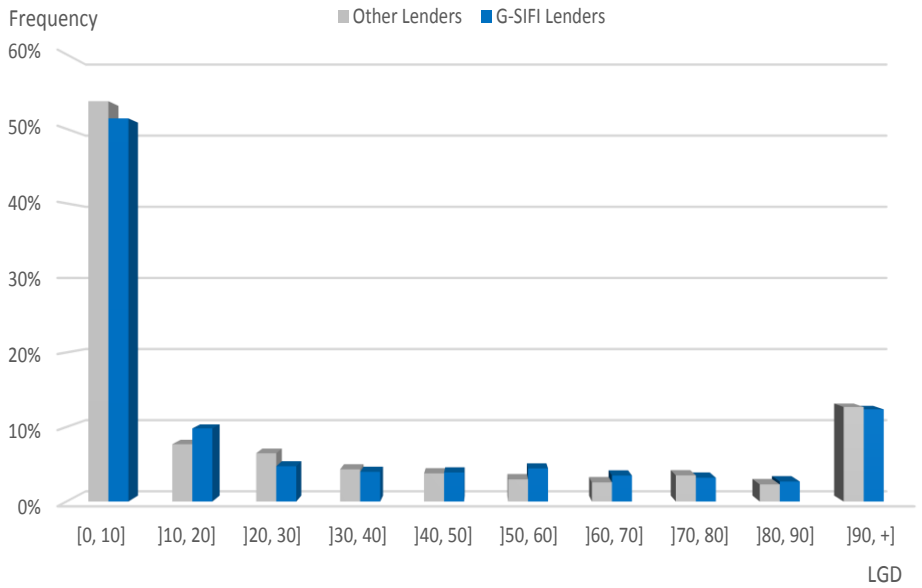

*Asset Class: Large Corporate unsecured, Europe*

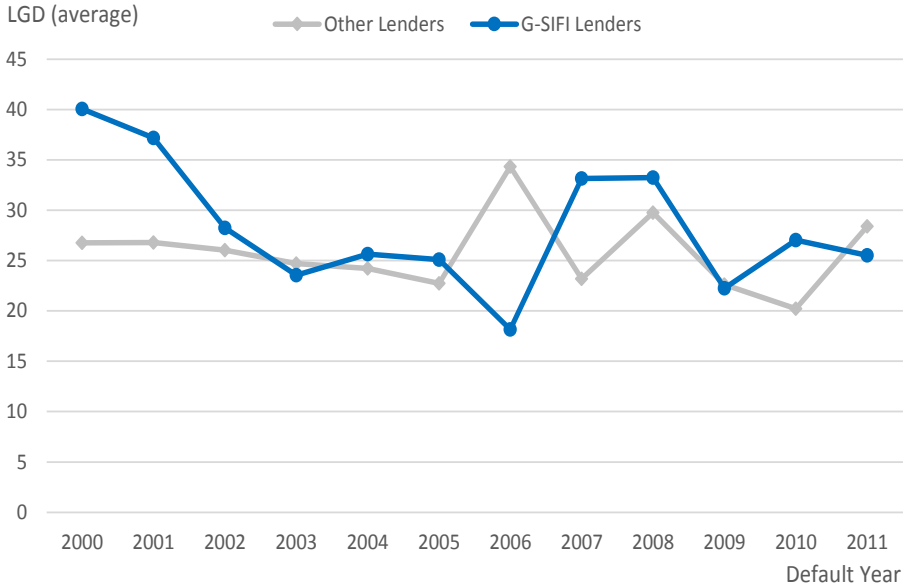




Global Credit Data

# G-SIFI Banks vs. Others: LGD




- Usual Bimodal LGD distribution
- no significant difference observed

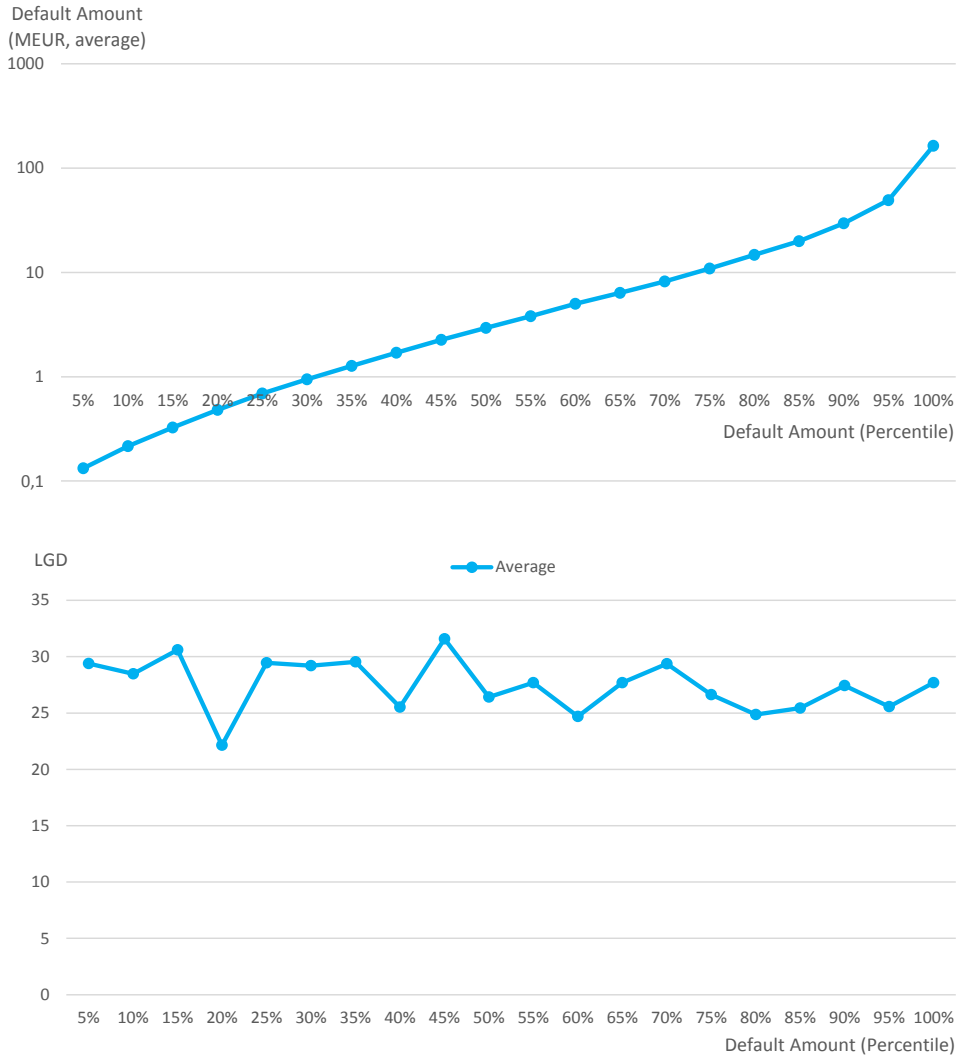
Table 4: G-SIFI vs. other lenders / LGD average and standard deviation over 12 years

| | Average | Standard Deviation |
|---|---|---|
| G-SIFI banks | 28.0% | 34.9% |
| Non G-SIFI banks | 25.5% | 35.4% |
| Combined | 27.5% | 35.3% |



Global Credit Data

# Relationship between the Exposure at Default (size) and LGD




- LGD appears constant across Exposure buckets
- Slightly decreasing (but remaining within the band 25-30% LGD)
- No significant/evident difference observed

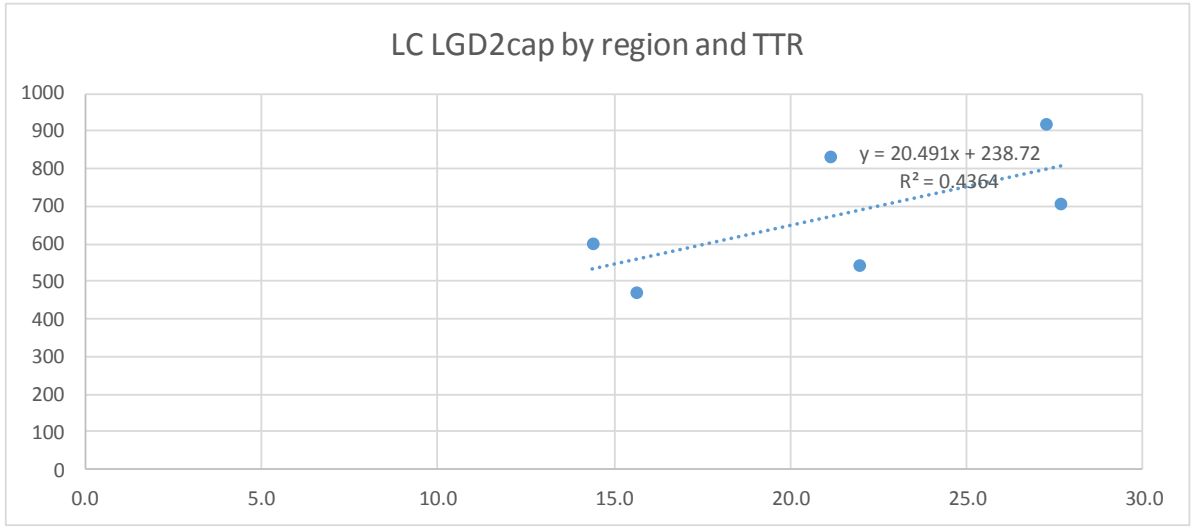


Global Credit Data

# Regional differences in LGD and TTR (time to resolution)

| Region | Average LGD% | Average time to resolution (days) | Number of Borrowers |
|---|---|---|---|
| East EU | 27.2 | 917 | 297 |
| Germany & Austria | 14.4 | 598 | 418 |
| Nordics | 15.6 | 469 | 531 |
| South EU | 22.0 | 539 | 279 |
| UK & Ireland | 27.7 | 703 | 616 |
| West EU | 21.2 | 830 | 816 |
| **Grand Total** | **21.3** | **687** | **2957** |

- workout timing differences between regions are fairly consistent over time
- Longer workout regions seem to correlate to higher LGDs, even at moderate discount rate (risk free)
- This may not be cause and effect






Global Credit Data

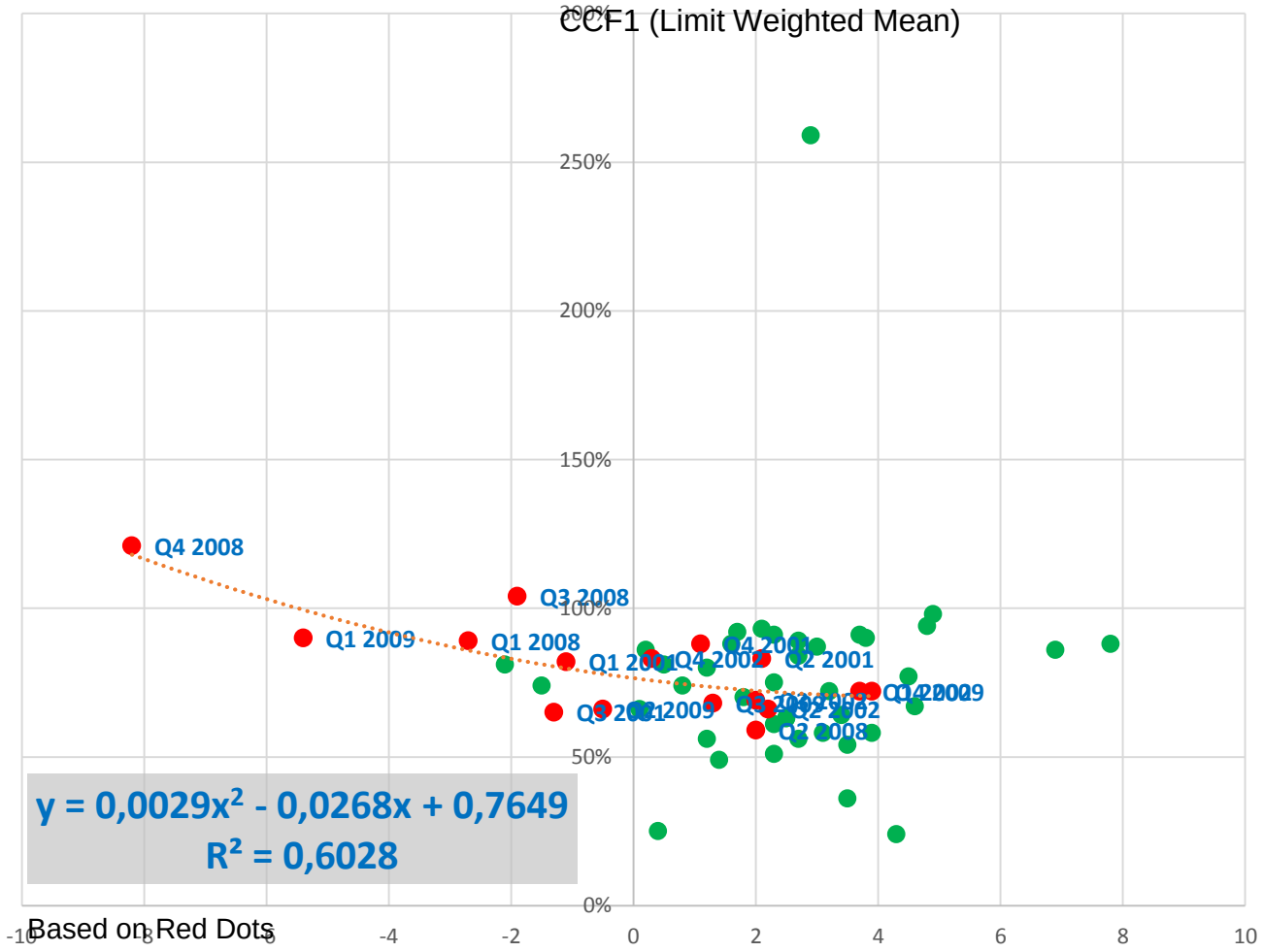
# Out of region Lending: TTR & LGD

| TTR | EU (B) | NA (B) | ROW (B) | Average |
|---|---|---|---|---|
| EU (L) | 2 years and 5 months | 2 years and 3 months | 2 years and 11 months | 2 years and 6 months |
| NA (L) | 2 years and 10 months | 2 years and 7 months | 2 years and 3 months | 2 years and 5 months |
| ROW (L) | 2 years | 2 years and 2 months | 2 years | 2 years |
| Average | 2 years and 6 months | 2 years and 4 months | 2 years and 5 months | 2 years and 5 months |

| LGD | EU (B) | NA (B) | ROW (B) | Average |
|---|---|---|---|---|
| EU (L) | 22,5 | 33,9 | 32,1 | 25,9 |
| NA (L) | 28,8 | 22,6 | 30,4 | 28,1 |
| ROW (L) | 25,2 | 30,0 | 44,1 | 36,9 |
| Average | 23,3 | 28,1 | 32,4 | 27,5 |

- In region lending: lower LGDs
- Out of region lending: higher LGDs

# CCF1 Stats Economic Cycle vs. CCF1, Time Series with a Lag




- Plot of GDP vs CCF1 with a Lag of 2 quarters: i.e. the CCF of quarter 'Q' is defined by the GDP Growth Rate of quarter 'Q-2'
- Some relationship between CCF and Economic Cycle
- worse the recession, higher the CCF

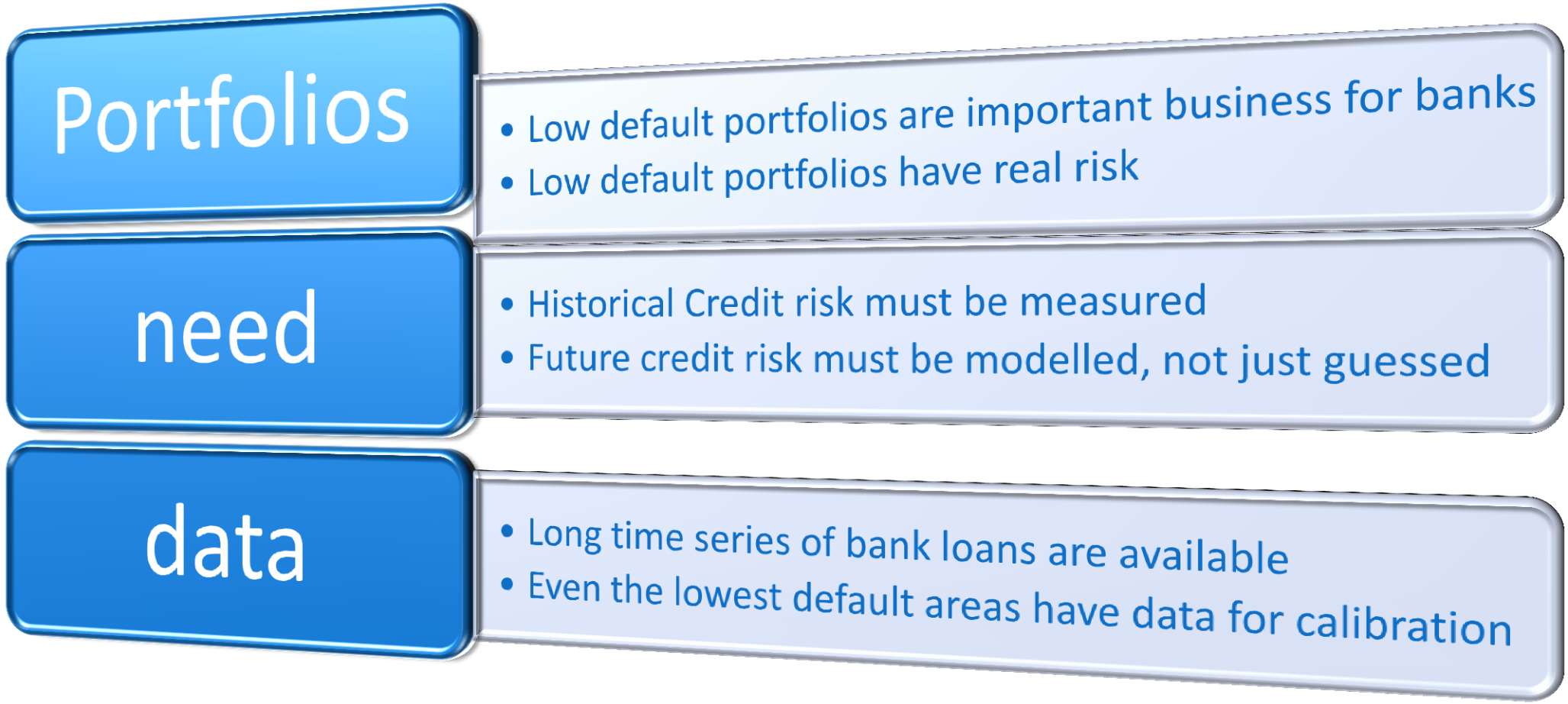


Global Credit Data

# Contents

- About Global Credit Data and why we are interested
- IRB constraints proposed by Basle committee
- How much data is required?
- How much data is available?
- Examples of use of available GCD Data
- Conclusion

# Conclusions





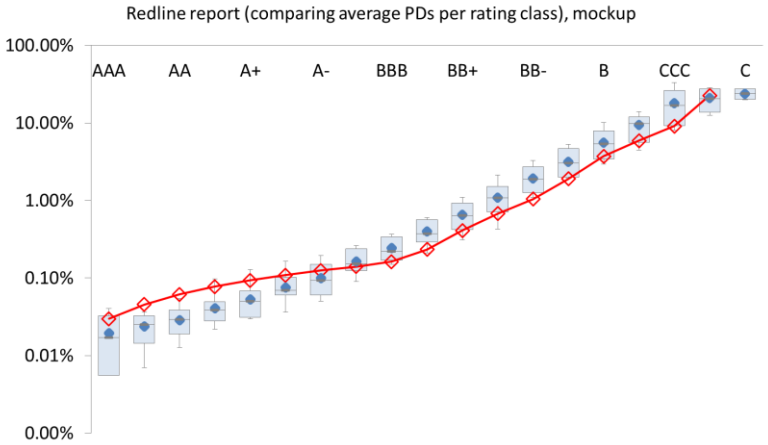

Global Credit Data

# Appendix

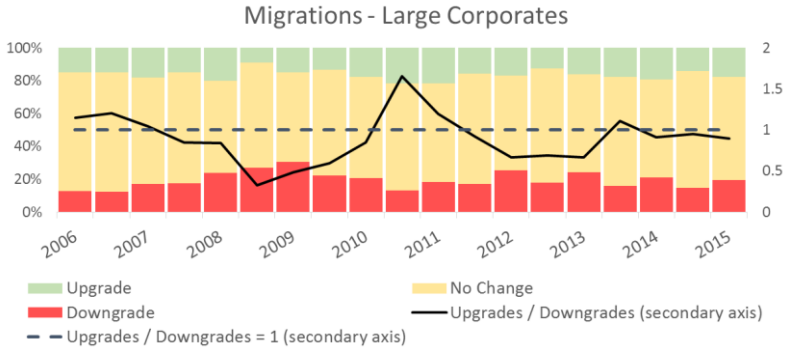
# Appendix: Usage of the PD ODF database

## A peek view into another GCD database

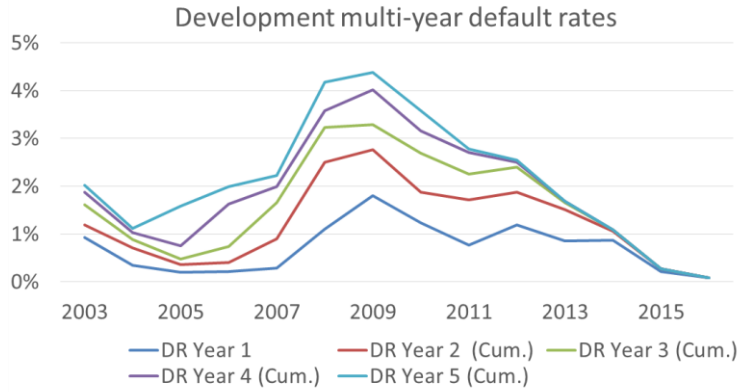
RegCapital: Benchmark PD masterscale vs peers




IFRS 9 / CECL: Extract a "systemic factor" from rating migrations or default rates




IFRS 9 / CECL: Benchmark your modelled PD term structures




RegCapital: Benchmark asset correlations and long term default rates

IFRS 9 – "Significant increase in credit risk": Determine thresholds for your "allowed" variance in lifetime PD

Overall: Benchmark your PD modelling techniques

# Appendix: GCD Member Banks April 2017

## Europe

ABN-AMRO
Allied Irish Bank
Banco Santander
Barclays
BNP-Paribas
Credit Agricole CIB
Credit Suisse Securities (Europe)
HSBC Trinkaus & Burkhardt
ING Bank
KfW Bankengruppe
Natixis
Nibc Bank
Raiffeisen Bank International
Rabobank
Royal Bank Of Scotland
Societe Generale
UBS
Unicredit S.p.A.

## Nordic

Danske Bank
DNB Bank
Nordea
Nordic Investment Bank
Nykredit
Swedish Export Credit
Skandinaviska Enskilda Banken
Svenska Handelsbanken
Swedbank

## Asia Pacific:

ANZ
Bank of East Asia
Commonwealth Bank Of Australia
National Australia Bank
Westpac

## Africa

ABSA
First Rand Bank
Investec Bank
Standard Bank of South Africa

## North America

Bank of America
Bank of Montreal
Bank of Nova Scotia
Capital One
CIBC
Citibank
Comerica
Desjardins
Export Development Canada
J.P. Morgan
M&T Bank
National Bank Of Canada
Northern Trust
PNC
Royal Bank of Canada
MUFG Union Bank
Wells Fargo