# LGD Parameter Estimation for the Retail Mortgage Book Based on EBA Guidelines

**Naval Chaudhary, Manish Malik**

## Abstract

EBA has published Guidelines (EBA/GL/2017/16) on PD, LGD Estimation and the treatment of defaulted exposures in April 2018. All the banks either have their business or headquarters in Europe need to implement these guidelines by end of 2020 and these could soon be adopted by other regulators. Purpose of this paper is to share a simplified interpretation of the guidelines from LGD perspective. Retail Mortgage portfolio has been used as an example to explain the concepts behind detailed guidelines in GL related to portfolios secured with collateral.
This paper covers a detailed view of LGD estimation process as per the guidelines. In this paper, a top down approach is adopted to bring together all the steps, procedures and analysis that form part of LGD parameter estimation.

## Keywords

PD, LGD, EAD, EBA, Mortgage, Quantification, Long Run Average, Downturn, Outcome, Resolved, Unresolved, Outcome window, Maximum Recovery Window

## 1. Introduction

Since last few years, regulators have been collecting data from Banks in the form of QIS to understand the variability in capital numbers reported by the banks. EBA issued detailed consultative guidelines (EBA/CP/2016/21) in November 2016 to reduce the unjustified variability. These guidelines were focused on the definitions and modeling techniques used in the estimation of risk parameters which were then published as EBA/GL/2017/16 in April 2018 after analyzing the responses submitted by various banks. These guidelines need to be abide by the banks who either have their business or headquarters in Europe. Based on the relationship between EBA and BOE, these guidelines are expected to be released as PRA guidelines with some minor adjustments.

These guidelines are very detailed and sometimes become too complex to understand the actionable. This paper unfolds the detailed guidelines into simplified process to compute LGD by adhering to the guidelines. All the details mentioned in this paper are author's personal views and are not related to their employer or any other 3rd party.

For the sake of consistency, terminologies are used in same context and meaning as referred by the GL. Tone and language of this paper is kept as simple as possible, in order to facilitate better and deeper understanding of concepts. In the paper, you may find a number of tables/illustrations have been used at various stages. Numbers mentioned in these tables are all random/fictional numbers and have been created by author just to facilitate understanding of discussed concept.

Due to the detailed guidelines, this paper has a lot of information on LGD and it might need multiple readings to fully understand the concepts explained. Group discussions might even be a better idea. Let's begin, Shall we?

## 2. Background

Before, we look into detailed story; let's look at below breakdown of process:

RWA (aka regulatory capital)

Banks keep aside provisions for Expected Losses (EL). UL or Unexpected loss is a measure of losses that bank might have in a stressed macroeconomic environment. Regulators want banks to keep aside UL as capital as a step towards future preparedness.

$$K = LGD \cdot \left[ N\left( \sqrt{\frac{1}{1-R}} \cdot G(PD) + \sqrt{\frac{R}{1-R}} \cdot G(0.999) \right) - PD \right]$$

$$RWA = K \cdot 12.5 \cdot EAD$$

For defaulted assets, EL is rather referred to as ELBE (Expected Loss Best Estimate). This concept is not covered in this paper

Our discussion in this paper is limited to LGD, which is conceptually the most complex parameter. In order to calculate capital we need LGD values, but single LGD value is not used for the whole book; rather in AIRB approach LGD is calculated at a pool level. Refer to table

| LGD Quantification/LGD Value | LGD Pools/LGD Modeling |
|---|---|
| In order to arrive at pool level LGD value, historic observed monthly LGD timeseries is used. LGD is max of Downturn LGD and Long-run Average LGD over a length of data, usually more than 5 years | Arrived at by using various statistical techniques/business judgement<br><br>Separate pools for Observation Defaults and Non Defaults |
| How to arrive at monthly LGD values? Use account level LGDs of that month.<br>What should be the length of historic data? Minimum 5 years recommended. | What should be the length of historic data? Minimum 5 years recommended. |

**LGD Computation/Account level LGD calculation process**

What population to be used for LGD calculation?
Different criteria for observation default and non-default

A particular month's LGD for Observation defaults will be based on observation defaults part of that pool for that month. If there are 2000 default accounts, their actual/estimated LGDs will be used to calculate pool level LGD

In case of Non defaults, the accounts from that pool which were non default as of Jul14 (for example) but defaulted in the next 12 months of performance; their actual/estimated LGDs will be used for calculating non default pool level LGDs

Based on the above display, there are 3 major stages of LGD parameter calculation:

1. LGD Computation – The analysis and processes that form part of the actual account level LGD calculation
2. LGD Modeling – Modeling/segmentation to create homogenous Risk Grades/pools using actual account level LGDs
3. LGD Quantification – Using historic account level LGDs over more than 5 years of data to arrive at pool level LGDs to be implemented for RWA calculation (Both Long run & Downturn LGD calculation)

Before we dive into details of LGD calculation methodology, let's assume we have a bank XYZ, with a mortgage book (for the sake of simplicity, let's assume it's purely first lien closed end). As of Dec'18, this mortgage book has 100,000 accounts. Out of these 100,000 accounts, 10,000 are defaults at observation and remaining 90,000 are non-defaults (For definition of Default refer to EBA/GL/2016/07). Let's assume that each of the account owes us $100,000. That means, together these 100,000 accounts owe $10bn to the bank. Now our job is to estimate appropriate capital (RWA). This bank uses AIRB approach to estimate its capital for this mortgage book. As part of the AIRB approach, an institution needs to estimate PD, EAD and LGD parameters at risk grade level which are used as an input to calculate RWA (capital). Below illustration will help to understand implementation view.

**Table 1.** Rating System Implementation view (numbers in millions)

| Segment | PD Pool | LGD Pool | Volume | PD | EAD | LGD | UL | EL | RWA |
|---|---|---|---|---|---|---|---|---|---|
| Non- | Low | Low | 10,000 | 2% | $ 1,000 | 10% | $ 16 | $ 2 | $ 195 |
| Default | Low | Med | 10,000 | 2% | $ 1,000 | 30% | $ 47 | $ 6 | $ 586 |
| | Low | High | 10,000 | 2% | $ 1,000 | 60% | $ 94 | $ 12 | $ 1,172 |
| | Med | Low | 10,000 | 5% | $ 1,000 | 10% | $ 26 | $ 5 | $ 329 |
| | Med | Med | 10,000 | 5% | $ 1,000 | 30% | $ 79 | $ 15 | $ 988 |
| | Med | High | 10,000 | 5% | $ 1,000 | 60% | $ 158 | $ 30 | $ 1,976 |
| | High | Low | 10,000 | 10% | $ 1,000 | 10% | $ 36 | $ 10 | $ 454 |
| | High | Med | 10,000 | 10% | $ 1,000 | 30% | $ 109 | $ 30 | $ 1,363 |
| | High | High | 10,000 | 10% | $ 1,000 | 60% | $ 218 | $ 60 | $ 2,725 |
| Sub total | | Non-Defaults | | | $ 9,000 | | $ 783 | $ 170 | $ 9,790 |
| Default | - | Low | 7,000 | 100% | $ 700 | 25% | ?? | $ 175 | |
| | - | High | 3,000 | 100% | $ 300 | 55% | ?? | $ 165 | |
| Sub Total | | Defaults | | | $ 1,000 | | ?? | $ 340 | |
| Total | | Whole Book | | | $ 10,000 | | ?? | $ 510 | |

Based on above capital calculation illustration, we need to answer 2 questions to be able to calculate RWA (from an LGD perspective). These are:

1. How did we arrive at these LOW, MED, HIGH LGD segments? The process of arriving at these LOW, MED, HIGH LGD segments is known as LGD Model development. One needs to use actual account level LGDs and a predictive modeling technique like decision tree or linear regression (segmentation of predicted LGDs) to arrive at this segmentation. Some of the key drivers for LGD could be LTV (Loan to value ratio), updated property value, macroeconomic variables etc.
2. How did we come up with those percentage LGD values which are used in calculating RWA?

As already mentioned, the process of arriving at final LGD values used for RWA calculation is known as LGD Calibration or LGD Quantification. In this paper, LGD Quantification is the term that has been used more predominately.

Before we look at the illustration of LGD quantification, let's talk about some ground rules set by regulatory guidelines for this process:

1. A given book should be segmented to homogeneous risk grades (LGD Pools).
2. Pool level LGDs are arrived at by calculating max of long run average and downturn LGD. Long run average simply means taking defaulted weighted average of historic monthly LGDs of mortgage book at LGD pool level. Downturn LGD refers to average LGDs pertaining to observed downturn period in Bank's data. If actual LGD data for downturn period is not available, LGD needs to be extrapolated. Methodology related to arriving at downturn period and the computation of downturn LGDs in not covered in this paper. Details on downturn LGD can be found in EBA/GL/2019/03 and EBA /RTS/2018/04.
3. Use of actual/realized LGDs to arrive at LGD parameters. Let's understand this remark in more detail. The objective of LGD model development process is to provide for risk differentiation by creating risk grades/segments/pools, and not to predict account level LGDs. This statement is somewhat debatable as this has not been explicitly called out in the GL, but it is our derived interpretation based on detailed readings and understanding of the guidelines. So, the purpose of model development is to only assign historical default accounts as well as current book (default & non-default) into LOW, MED, HIGH LGD pools (Use of 3 pools is just for illustration). Average actual LGDs of the accounts that form part of LOW, MED, HIGH LGD pools will be used to arrive at final pool Level LGD parameter by computing maximum of long run average and downturn LGD, as explained above. The use of Long run averages is quintessential condition to the LGD Risk parameter estimation.

The below illustration gives us a glimpse of Long run average LGD calculation for LOW LGD segment for both Observation Default and Non Default population:

**Table 2.** Long Run Average Calculation.

| Month | Observation Defaults Low LGD Segment | | Observation Non Defaults Low LGD Segment | | |
|---|---|---|---|---|---|
| | # Defs | LGD | # Obs NDs | # Perf Defs | LGD |
| Jan-07 | 2,513 | 37.50% | 20,454 | 202 | 23.40% |
| Feb-07 | 2,789 | 42.10% | 21,300 | 238 | 17.90% |
| Mar-07 | 2,654 | 35.70% | 22,456 | 196 | 18.60% |
| Apr-07 | 3,012 | 34.60% | 23,423 | 199 | 19.50% |
| May-07 | 3,208 | 44.10% | 23,999 | 247 | 27.00% |
| Jun-07 | 2,987 | 28.70% | 25,011 | 226 | 22.40% |
| - | - | - | - | - | - |
| - | - | - | - | - | - |
| - | - | - | - | - | - |
| - | - | - | - | - | - |
| Sep-17 | 5,899 | 47.90% | 45,432 | 501 | 32.00% |
| Oct-17 | 5,975 | 43.80% | 47,623 | 537 | 28.70% |
| Nov-17 | 5,701 | 38.40% | 49,754 | 498 | 25.20% |
| Dec-17 | 6,003 | 39% | 51,010 | 568 | 22.40% |
| **Final Pool level Long-run LGD** | | **40.1%** | | | **24.9%** |

This illustration assumes that XYZ bank has actual default and recovery data available in their data marts starting from Jan 2007. Not all the organisation might have this much length of data available. In that scenario, organisations should refer to guidelines specific to historical observation data and related characteristics in GL.

Let's look at table 2 more closely now. For Observation Non-Defaults, there were 21,300 Non defaults (as of Feb2007) in Low LGD segment, out of which 238 of them defaulted in the next 12months. When these 238 defaulted accounts were tracked further in future for recoveries/losses, they together had a representative LGD of 17.9%. LGD stands for Loss given default. For accounts that do not default, you don't know what loss they could have given. But it is only a logical derivation that, if they would have defaulted, their loss rates would be similar to the ones who did. One cannot come up with actual LGD for all the 21,300 accounts, but one can safely assume that if all of them were to default, their LGDs would look similar to LGDs of those 238 accounts.

Let's come to Observation defaults. Similar to above explanation, the interpretation goes like this. There were 2,789 accounts that were part of the LOW LGD Observa-

tion default segment as of FEB 2007 snapshot. Now, one key thing to visualize here is that these 2,789 accounts would have served varying months in default, while they all sit together in the observation default segment. These are all non-closed, non-charged off accounts. Some of them could have defaulted 5 years ago, while others could have defaulted recently (<6 months ago). Naturally, it might seem unfair to give them similar LGD treatment. Their LGD calculation and segmentation is little complex than observation non defaults. The details of their differential treatment are covered as part of the Reference Date discussion.

The way to arrive at collective monthly LGD of 238 performance defaults (17.9% - shown in table 2) is to do EAD weighting of individual account level LGDs. While arriving at 40.3% overall LGD, we perform default weighting on monthly LGD values.

Just to highlight here, model development data preparation for Observation Default and Non-Default segmentation, also follows same principle of population definition as for LGD quantification. i.e. in order to use Jan 2010 month as part of model development, we will use 238 performance defaults and we will use their attributes available as of Jan 2010 (Non default snapshot) to explain the LGDs of these 238 accounts. Similarly, for observation default modeling, accounts that form part of Jan 2010 vintage as observation default should be used for model development. One may notice that now the construct of non-default LGD model looks like similar to PD model, with accounts duplicating in monthly data for more than one month. This is not considered as a challenge from quantification perspective, but for modeling exercise analyst may choose some appropriate de-duplication strategy. Details specific to observation default have been discussed later.

## 3. Account Level LGD Calculation

The most important aspect of this discussion is how to calculate account level LGDs (aka LGD Computation). We will proceed forward with an example of observation non-default. The principles used for observation default LGD calculation are same as explained for non-defaults. The only exception lies with respect to use of reference dates and incomplete recovery processes. They have been explained in detail towards the end of document.

What is LGD?
LGD is loss given default. It is a percentage estimate of economic loss given an EAD value.

Loss= EAD – Recovery + Costs Incurred

LGD = Loss / EAD

We will understand the account level LGD computation and associated analysis/principles with the example of 9 accounts from Feb 2007 snapshot as shown table 3. Since, we are looking at LGD from an observation Non default perspective; all of these 9 accounts were non-default as of Feb 2007 but defaulted sometime in between Mar2007 to Feb 2008. Post default these accounts have been tracked for recoveries/losses/costs-incurred.

**Table 3.** Default account recovery performance

| | Default Month | Aug-07 | Jun-07 | Oct-07 | Apr-07 | Jul-07 | Aug-07 | May-07 | Jan-08 | Oct-07 |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Month** | **Paid Full** | **$0 Recovery Write Off** | **Some Recory Write Off** | **REO Sold** | **REO** | **Cured Re-Default** | **Cured** | **Unresolved** | **??** |
| | | **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** | **I** |
| Fixed 12 month performance window for observing default event | Feb-07 | 75,000 | 75,000 | 75,000 | 75,000 | 75,000 | 75,000 | 75,000 | 75,000 | 75,000 |
| | Mar-07 | | | | | | | | | |
| | Apr-07 | | | | EAD : 70,000 | | | | | |
| | May-07 | | | | | | | EAD : 70,000 | | |
| | Jun-07 | | EAD : 70,000 | | | | | | | |
| | Jul-07 | | | | | EAD : 70,000 | | | | |
| | Aug-07 | EAD : 70,000 | | | | | EAD : 70,000 | | | |
| | Sep-07 | | | | | | | | | |
| | Oct-07 | | | EAD : 70,000 | | | | | | EAD : 70,000 |
| | Nov-07 | | | | | | | | | |
| | Dec-07 | | | | | | | | | |
| | Jan-08 | | | | Rec: 1,000 | | | | EAD : 70,000 | |
| | Feb-08 | | | | | | | | | |
| | Mar-08 | | | Rec: 2,000 | | | | | | |
| | Apr-08 | | | | | | | | | |
| | May-08 | | | | | | Rec: 500 | | | |
| | Jun-08 | | | Rec: 5,000 | | Rec: 5,000 | Rec: 500 | | | |
| | Jul-08 | | | | | | Rec: 500 | Rec: 500 | | |
| | Aug-08 | Rec: 70,000 | | | | | Rec: 500 | Rec: 500 | | |
| | Sep-08 | | | | | | Rec: 500 | Rec: 500 | | |
| | Oct-08 | | | | Rec: 4,000 | | Rec: 500 | Rec: 500 | | |
| | Nov-08 | | | | | | Cured | Cured | | Rec: 1,000 |
| | Dec-08 | | | | | | Rec: 500 | | | Rec: 1,000 |
| | Jan-09 | | | Rec: 3,000 | | | Rec: 500 | | | |
| | Feb-09 | | | | | | Rec: 500 | | | |
| | Mar-09 | | | | | | Rec: 500 | Re - Default | | |
| | - | - | - | - | - | - | - | - | - | - |
| | - | - | - | - | - | - | - | - | - | - |
| | Jan-10 | | | | | | Re - Default | | | |
| | Feb-10 | | | | | | | | | |
| | Mar-10 | | | Rec:10,000 C/O: 50,000 | | | | Rec: 10,000 | | Rec: 2,000 |
| | Apr-10 | | | | | | | | | Rec: |

| Default Month | Aug-07 | Jun-07 | Oct-07 | Apr-07 | Jul-07 | Aug-07 | May-07 | Jan-08 | Oct-07 |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | 1,000 |
| May-10 | | | | | | | | | Rec: 1,000 |
| Jun-10 | | | | | | | | | |
| Jul-10 | | | | | | | Rec: 33,000 | | |
| Aug-10 | | C/O: 70,000 | | | | | | | |
| Sep-10 | | | | | | | | | Rec: 1,000 |
| Oct-10 | | | | | | | | | Rec: 1,000 |
| Nov-10 | | | | | | | C/O: 25,000 | | |
| Dec-10 | | | | REO (Book value: 55,000) | | | | | |
| Jan-11 | | | | | | | | | |
| Feb-11 | | | | | | | | | |
| Mar-11 | | | | | | | | | Rec: 2,000 |
| Apr-11 | | | | | | | | | |
| - | - | - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - | - | - |
| Jan-12 | | | | | REO (Book Value: 40,000) | | | | |
| Feb-12 | | | | | | C/O: 40,000 | | | |
| Mar-12 | | | | REO Sold (50,000) | | | | | |
| Apr-12 | | | | | | | | | |
| May-12 | | | | | | | | | |
| Jun-12 | | | | | | | | | |
| Jul-12 | | | | | | | | | |
| Aug-12 | | | | | | | | | |
| Sep-12 | | | | | | | | | |
| Oct-12 | | | | | | | | | |
| Nov-12 | | | | | | | | | |
| Dec-12 | | | | | | | | | Rec: 45,000 |
| - | - | - | - | - | - | - | - | - | - |
| - | - | - | - | - | - | - | - | - | - |
| Oct-17 | | | | | | | | | |
| Nov-17 | | | | | | | | | |
| Dec-17 | | | | | Never Sold | | | Still open and never cured | |

Some numbers provided in the above table might seem unrealistic but they have been purposefully kept that way in order to simplify things.

In the above paragraph, while we mentioned that we track accounts post default for recoveries/losses; someone might ask, how long we should keep tracking an account. There has to be a stopping point. This is where we introduce the concept of Maximum Recovery Window. As per GL

"*156. Institutions should define the maximum period of the recovery process for a given type of exposures from the moment of default that reflects the expected period of time observed on the closed recovery processes during which the institution realises the vast majority of the recoveries, without taking into account the outlier observations with significantly longer recovery processes. The maximum period of the recovery processes should be specified in a way that ensures sufficient data for the estimation of the recoveries within this period for the incomplete recovery processes.*"

Don't worry about what incomplete recovery processes mean yet. Key takeaways:

1. Max recovery window is the duration (in months/years) for which an account should be tracked for recoveries/costs post default.

2. This window length should be defined long enough to realise vast majority of recoveries with only outliers sitting outside that window.

3. Should be based on institution's own experience.

What analysis can be conducted to find maximum recovery window length?

We have shared our thoughts later in this document. For now, let's assume it is 5 years (60 months) for our example.

There is another term that has been mentioned above is 'Closed Recovery Process'. Closed recovery process means a Resolved account. An account where you have an outcome/event that leads to closure of story from recovery stand point. Below are some of the key closure outcomes (resolve categories) used in the paper:

1. Paid in Full – The customer chooses to close the account by paying full outstanding amount.

2. Partial/Full Charge off taken (Write Off) – Bank is able to recover either some amount of money or nothing, before bank decides there is no point chasing the account in court for collateral (since we have taken mortgage as an example). So bank writes off this asset (account) from the book. Just to be mindful, sometimes bank takes partial charge off without closing the account. Closure of loan is important for resolved classification.

3. Collateral owned by bank – Post default event, Bank claims its right to ownership of collateral in a court. Post winning the litigation, property is legally owned by bank and customer is freed from the financial obligations to bank. Bank later on sells the property to recover cash.

4. Cured – Customer comes out of his financial troubles and starts to pay back. This can happen with or without modification/restructuring of loan. A customer is flagged as default by certain number of triggers. While some of them can be curable, others may not be. Triggers like delinquency (90+ DPD) are softer and more likely to cure. However regulators don't want to call an account non-default as soon as an account

goes below 90 DPD. Instead, we should wait for it to stay out of all the default triggers for certain number of months. This duration is known as probation period. An analysis can be conducted to come up with an optimal probation period for each curable trigger. So, Probation period is additional time required to be spent in default category (even post receiving continuous payments/coming out of default condition), before account can be considered as cured. Customer cures back to non-default status after serving appropriate probation period. But this story doesn't end here. We will talk about independent and dependent re-default events for cured accounts in a bit. Let's assume probation period to be 6 months for all curable default triggers. For more details regarding probation period, please refer to section 7 of EBA/GL/2016/07 on Default definition.

Now, if an account is not able to reach any of the above stated stage within the Maximum recovery window, it is called an Unresolved account.

But what about those accounts who have recently defaulted and did not get 60 months (max recovery window) for closure (resolution). These accounts are called as Incomplete recovery processes. They will be discussed, when we go back to LGD quantification piece. For now, let's focus back on LGD computation.

In all this chit chat, we almost forgot about the table 3. Let's attend to that now. We will talk about the story of each of the 9 accounts in detail. Here we go:

**Account A:** This account had a balance of $75,000 as of Feb2007. This customer made some payments in between, but finally triggered default in the month of Aug with an outstanding balance of $70,000. This includes unpaid fees and interest. For more detail on regulations related to treatment of unpaid fees and interest, please refer to EBA/GL/2017/16 (Section 6.3.1.2). Post default, collection's team gets in touch with this customer and makes all possible efforts to bring customer back to making payments or recover something. Customer was able to manage a refinance deal with some other bank. In Aug 2008, other bank pays full outstanding on behalf of the customer and we close the loan. So, this customer is tagged as 'Paid in Full' (Resolved).

**Account B:** This account also defaults with $70,000 in the month of Jun 2007. Collection's team worked really hard on this account but hard luck. The customer's house, that he had kept as collateral, has depreciated to <$30,000. The collection's manager thinks that it is not worthy to spend money on lawyer fees, property taxes etc. to recover from this account. Eventually, we take a full write off of $70,000 and close this account.

**Account C:** This account defaults in the month of Oct 2007. Collection's team is able to recover a sum total of $ 10,000 from this account, but eventually had to write off the remaining balance without any more success. This customer is flagged as 'Partial Charge off' (Resolved).

**Account D:** This account defaults in the month of Apr 2007. Collection's team is able to recover $5,000 from this customer, before we win litigation against this customer in court and take over his underlying collateral. This event is also known as Real Estate owned (REO). We evaluate the market value of property as of Dec 2010 (month of REO). We close the customer's account from our books, but book another asset (real estate) on our ledger with a value of $55,000. In this whole transaction, bank lost a net value of $10,000 (70,000-5000-55000). Bank takes a write off equivalent to

$10,000. Later, bank is able to sell the same property for $50,000 in the month of Mar 2012 with an additional write off of $5,000. What should be the Loss for such an account? Ideally, it should be $15000. But as per regulatory guidelines, the closure of this account happened at the time of REO and not at REO sale (very valid point). So, should we assume the recovery as the property valuation at REO with an adjustment factor (haircut) to accommodate for losses at the time of sale? We will take this up in more detail when we talk about REO Haircut analysis. For more details refer to EBA/GL/2017/16 (Section 6.1.3).

**Account E:** This account defaults in the month of Jul 2007. We are able to recover $5,000 of cash before we do an REO for this account in Jan 2012. The market value of account is $40,000 at the time of REO. We haven't been successful to sell this property as of end of Dec 2017 and it still sits on our books as an asset. We can use adjustment factor calculated using REO Haircut analysis (explained later) to come up with final recovery value.

**Account F:** This account defaults in the month of Aug 2007. Collection's team is able to work out a repayment plan. Customer starts paying us back at a rate of $500 per month starting May 2008. With a probation period of 6months, we eventually cure this account back to non-default status as of Nov 2008. This continues to pay us for some more months, before re-defaulting in the month of Jan 2010. This account is later on charged off without any further recoveries in the month of Feb 2012. What should be the loss for this account? Should we call it cured or not? Let's talk about independent vs. dependent default events.

In an ideal scenario a customer defaults (misses on their financial obligation to pay back the bank), only when there is a financial trouble in borrower's life. Now whatever that might be, a customer can come out of it and may start paying us back. The ideology here is that, if a customer manages to stay out of default (post curing) for 9 months in continuity, then any default event that happens post that is assumed to be have been triggered by an entirely new trouble; and thus is an independent default event. However, if the default event reoccurs within the 9 month window from curing, it is considered to be a connected/dependent default event and is considered to be the continuity of the same trouble he /she had faced earlier. This 9 months additional requirement is one of the latest additions by EBA to regulatory requirements. This entire trouble story is just to simplify things. By the way, this 9 month of observation window is over and above probation period (total 9+6=15). So, in nut shell if an account re-defaults within 9 months of curing, that scenario is considered to be as a continuous default event starting from Aug 2007, without the consideration of curing in between. However, if this customer re-defaults post 9 months of window that is considered as an independent default event and customer is considered to have cured. In case of account F, it is considered to have cured as of Nov 2008. This assumption will not lead to under estimation of LGD, as this account will come back in LGD data as of month Jan 2010 and will be able to contribute its true LGD. Couple of things to mention here are that as per regulatory guidelines:

1. For a cured account; all the outstanding balance as of the month of cure is assumed to have been recovered in full. However, discounting of cash flows still needs to be performed for LGD calculation sake.

2. This additional 9 month window is specific to LGD and should not interfere with default tagging from PD and capital calculation perspective.

**Account G:** We already covered the story for this account while taking about account F. This account defaults in the month of May 2007, cures in Nov 2008 but re-defaults in Mar 2009. We do not consider this account as cured, and rather continue to track its recovery further. We are able to recover in total $45,000 from this account, but eventually take a write off of $25,000 in Nov 2010. This account is considered to have resolved by 'Partial Charge Off' category.

**Account H:** This account defaulted in the month of Jan 2008. We haven't had any success with this account. As of Dec 2017, this account is neither Cured nor resolved by any other means. It is classified as Unresolved account. Since this is not a closed recovery process, that cannot be the justification to exclude this account. As per the regulatory guidelines, no default account should be excluded from LGD estimation, in any case. For more details refer to EBA/GL/2017/16 (Section 6.1.2 and 6.3.2.1).

But what should be the appropriate LGD for this account, we will talk about it when we discuss Unresolved LGD estimation. Let's park it aside for now.

**Account I:** This account defaulted in the month of Oct 2007. The account continues to make some inconsistent payments; until it finally pays us the total left over amount of $45,000 and becomes Paid in Full. A key thing to note here is that the final closure of account happened in the month of Dec 2012, which is out of the Maximum Recovery Window length (Nov2007 to Oct2012). What should we do with such an account? Is it resolved? Should we include his recoveries and costs observed up till Oct 2012 only and call it Unresolved?

What we are about to say is no-where directly mentioned in guidelines but it is our interpretation based on multiple readings of GL. At many places regulatory guidelines becomes subjective and open to reader's interpretations. We shall have no concerns, if somebody else has a different viewpoint on a same scenario provided it adheres to guidelines from GL. As per GL, for any defaulted account; if you can calculate their actual LGD based on data (using information even past Maximum recovery window), bank should use that actual realized LGD in the LGD quantification. Paragraph 156 and 157 in GL (EBA/GL/2017/16 - Section 6.3.2.3) can be referred as support to our opinion.

This brings us to the idea of never ending tracking of defaulted accounts until resolution and in case of cures even 9 months post month of cure. GL talks about including all the historic default accounts for LGD quantification. For the accounts that resolved irrespective of the Max recovery window length, we know their actual LGD (Account I being one such example). So for all such cases, the recommendation is to use their actual LGDs in the LGD quantification. For the ones that have not resolved, there are 2 possibilities:

> If an account has had performance available for a period longer than or equal to maximum recovery window; then we call it Unresolved and calculate LGD according to guidelines specific to such cases.
>
> Otherwise, we use methodology discussed for Incomplete recovery processes. Out of all the 9 examples discussed above, none of them corresponds to Incomplete recovery process. We will discuss it in a bit.

Below table, shows the LGD calculation for the 9 of above discussed examples.

Table 4: LGD Calculation for Accounts Covered in Scenario A to I

| Account | EAD | Recovery | Discounted Recovery | Direct Cost | Indirect Cost | Loss | LGD |
|---|---|---|---|---|---|---|---|
| **A** | $ 70,000 | $ 70,000 | $ 60,000 | $ 0 | $ 700 | $ 10,700 | 15% |
| **B** | $ 70,000 | $ 0 | $ 0 | $ 0 | $ 1,200 | $ 71,200 | 102% |
| **C** | $ 70,000 | $ 20,000 | $ 17,000 | $ 0 | $ 1,150 | $ 54,150 | 77% |
| **D** | $ 70,000 | $ 50,000 | $ 42,000 | $ 2,000 | $ 700 | $ 30,700 | 44% |
| **E** | $ 70,000 | $ 40,000 | $ 35,000 | $ 2,000 | $ 600 | $ 37,600 | 54% |
| **F** | $ 70,000 | $ 70,000 | $ 65,000 | $ 0 | $ 400 | $ 5,400 | 8% |
| **G** | $ 70,000 | $ 45,000 | $ 39,500 | $ 500 | $ 800 | $ 31,800 | 45% |
| **H** | $ 70,000 | ?? | ?? | ?? | $ 1,500 | | |
| **I** | $ 70,000 | $ 70,000 | $ 58,000 | $ 0 | $ 600 | $ 12,600 | 18% |

In the above table, you may notice recoveries have been discounted back to the month of default to calculate final recovery values to be used for LGD calculation. This is one of the regulatory requirements. Also, one may notice use of 2 different types of costs for LGD calculation. Direct cost is money spent by bank in recovery process which can be directly attributed to a particular account. Examples of direct cost can be Attorney fees, insurance, property tax payments etc. Indirect cost is something that cannot be directly attributed to a single costumer. One example can be administrative cost spent on maintaining collections team and systems. For more details on discounting and cost refer to Section 6.3.1.3, 6.3.1.4 of EBA/GL/2017/16. Another thing to note is that in above table actual notional recoveries have been used for LGD calculation instead of the one that will be estimated using analysis similar to explained below.

## 4. Supporting Analysis to LGD Computation

While discussing account level examples, we talked about few analyses like REO Haircut, Unresolved analysis etc. Let's now talk about framework for some of these analyses:

**Maximum Recovery Window Analysis:**

GL states in Para 156 of Section 6.3.2.3 that Maximum recovery window length should be long enough to exclude only outlier cases of recovery. Thus we propose to use a 99th percentile month value of all the recovery events. To put it in simple words, if there were 100 recovery events recorded from 20 defaults (every payment received post default is considered as a separate recovery event) and they are sorted in ascending order of month of payment from the default month; 99th value will the used as the length of Maximum recovery window. Another approach can be to plot cumulative % $ recovery by months spent in default and arrive at a saturation point to be used as Maximum recovery window. There can be many other ways to arrive at Max recovery window. Any different approach can be adopted to meet the specific scenarios pertaining to the portfolio under consideration, provided they comply with the underlying principles stated by GL.

**REO Haircut Analysis:**

As recommended by the GL, a default account should be assumed to have resolved at the moment of REO and not at the moment of REO Sale (Refer to Para 116,117 in Section 6.1.3 of GL). Institutions should use book value (Refer GL) of collateral at the time of REO and multiply it with a haircut factor to come up with adjusted recovery value. This new recovery value will now be discounted back to the month of default from the date of REO. Let's go back to our example of account D. What guidelines mean is that instead of using the actual $50,000 cash recovery, we should multiply $55,000 (book value at the time of REO) with a factor X, which captures 2 components:

- Any changes in property value post REO. In our case, that factor could be 50/55=0.91

- Capture the discounting effect of delayed reception of cash. In our case, while REO happened in Dec 2010, the property got sold only in Mar 2012 (15 months later). This factor will further reduce the recovery. Let's assume that factor is 0.87.

So, the estimated recovery at the time of REO for account D is now, $55,000 X 0.91 X 0.87=$43,543. This $43.5K is now discounted back to month of default from Dec2010 (month of REO). While we have represented the haircut analysis to be composed of 2 factors, but that was just for simplification purposes. The GL expects it to be a one single multiplier. However, there is no restriction to have same haircut for the whole book. One may choose to have different haircuts by LTV buckets, geographical location etc. with the support of rationale. One should use all the historic REO sold cases to compute this haircut factor. There are more detailed guidelines on this in the GL (Section 6.1.3) which talk about various different ways of calculating Haircut factor. But above mentioned example is considered sufficient to share the intent of the regulatory guidelines. We can use this REO haircut factor to calculate LGD of account E from the above Illustration.

**Unresolved Framework:**

As mentioned earlier, an unresolved account is one that doesn't get resolved even after performance of Maximum recovery window length. As per GL, there should not be a lot of such cases keeping in mind the definition of maximum recovery window. But what should be the LGD of these accounts?

As per GL, no estimation of future recoveries should be performed after the Maximum recovery window. Hence, we can only use recoveries and costs realized till the latest available snapshot to calculate their account level LGDs. This includes recoveries/costs realized post maximum recovery window. For support of this view refer to Paragraph 154, Section 6.3.2.3 and Paragraph 181, Section 7.3.1 in EBA/GL/2017/16. In Case of Mortgages (or other secured products), GL allows for the use of presence of collateral while estimating LGDs. So, one may use collateral value as of Max recovery window; overlay it with a conservative downward adjustment factor to consider it as recovery and estimate LGD. But for unsecured accounts, only actual recoveries and costs can be used to calculate LGDs, which may lead to very high LGD estimates for these accounts.

Now with REO Haircut analysis and Unresolved analysis in place we will be able to

assign LGD to every defaulted account. So if there were 238 performance defaults for the month of Feb 2007, we should be able to assign an LGD value to every account. We talked about aggregating account level LGDs to monthly LGDs by using EAD weighting.

However, what about accounts those were performance defaults of month Jan 2015. For them, we have performance available until Dec 2017. While it's possible that some of these defaulted accounts might have resolved, but for the remaining ones it will be unfair to consider all of them unresolved and give higher LGDs. Since these accounts didn't get time equal to maximum recovery window to show their true performance, they are called incomplete recovery processes. While some of us may think it to be a better idea to exclude such vintages, regulatory guidelines are more democratic in nature which intend to provide a right to representation in LGD for even young defaults. So, we will need another analysis to estimate the LGDs of this scenario. GL is very clear about not excluding any of the defaulted accounts, even If they are from recent vintages with insufficient performance available. Refer to discussion of Observed average LGD vs Long-run LGD on Pg 34 of GL. Below is our view of Incomplete recovery process analysis:

**Incomplete recovery process:**

In our viewpoint, the objective of this analysis should not be to directly predict actual LGD, but to estimate the proportion (probability) of different outcome categories as of end of max recovery window. In the beginning we had talked about 4 resolved categories and one unresolved category. The detailed illustration may help understand our view:

Let's assume the objective is to calculate LGD for Jan 2015 observation non defaults.

Table 5: Performance Defaults for Jan 2015 observation

| Non-Default Snapshot (Jan-15) | 20,000 |
|---|---|
| Performance Defaults (Feb-15 to Jan-16) | 210 |
| Resolved in available performance till Dec-17 | 45 |
| Incomplete Recovery Processes | 165 |

Based on table above, we should already have the actual LGD of 45 resolved defaults. But what should be the estimated LGD for remaining 165 defaults? These 165 defaults have months spent in default that range from 23 (Jan16 defaults) to 34 (Feb15 defaults).

Table6: Months in Default distribution

| Month of Default | Months Spent in Default till Dec-17 | Distribution |
|---|---|---|
| Jan-16 | 23 | 19 |
| Dec-15 | 24 | 12 |
| Nov-15 | 25 | 15 |

| | | |
|---|---|---|
| Oct-15 | 26 | 13 |
| Sep-15 | 27 | 16 |
| Aug-15 | 28 | 14 |
| Jul-15 | 29 | 8 |
| Jun-15 | 30 | 12 |
| May-15 | 31 | 14 |
| Apr-15 | 32 | 10 |
| Mar-15 | 33 | 15 |
| Feb-15 | 34 | 17 |

Further, these accounts can also be visualized to be belonging to different LTV buckets. LTV has been used for segmentation since it is one of the most important independent variables for LGD. However, any other suitable variable can also be used instead of LTV.

Table 7: Defaults Segmented by Months in Default and uLTV

| Months Spent in Default | uLTV Buckets | Volume | LGD |
|---|---|---|---|
| 23 | <30% | 9 | ?? |
| | 30-60% | 6 | |
| | >60% | 4 | |
| 24 | <30% | 5 | |
| | 30-60% | 4 | |
| | >60% | 3 | |
| 25 | <30% | 4 | |
| | 30-60% | 4 | |
| | >60% | 7 | |

The intent of this analysis is to come with an LGD value corresponding to months spent in default and LTV. So, by looking at these 2 aspects of an account, we can allocate a suitable LGD to the account.

Table below shows how it can be done.

Table 8: LGD Calculation by Months in Default and LTV Based Segment

| **Based on Full time series (performance till Max recovery window)** | | | | | |
|---|---|---|---|---|---|
| **Months in Default** | **uLTV Buckets** | **Outcome category** | **Distribution as of Max Recovery Window** | **Avg LGD** | **Final LGDs** |
| 24 | <30% | Paid in Full | 40% | 15% | **25%** |
| | | REO | 20% | 35% | |
| | | Charge Off | 10% | 60% | |
| | | Cure | 25% | 8% | |
| | | Unresolved | 5% | 75% | |
| | 30% - 60% | Paid in Full | 30% | 15% | **34%** |
| | | REO | 25% | 35% | |

| Based on Full time series (performance till Max recovery window) | | | | | |
|---|---|---|---|---|---|
| **Months in Default** | **uLTV Buck-ets** | **Outcome category** | **Distribution as of Max Recovery Window** | **Avg LGD** | **Final LGDs** |
| | | Charge Off | 20% | 60% | |
| | | Cure | 15% | 8% | |
| | | Unresolved | 10% | 75% | |
| | | Paid in Full | 10% | 15% | |
| | | REO | 40% | 35% | |
| | >60% | Charge Off | 30% | 60% | **42%** |
| | | Cure | 10% | 8% | |
| | | Unresolved | 10% | 75% | |
| | | Paid in Full | | | |
| | | REO | | | |
| | <30% | Charge Off | | | |
| | | Cure | | | |
| | | Unresolved | | | |
| | | Paid in Full | | | |
| | | REO | | | |
| 25 | 30% - 60% | Charge Off | Same as demonstrated above | | |
| | | Cure | | | |
| | | Unresolved | | | |
| | | Paid in Full | | | |
| | | REO | | | |
| | >60% | Charge Off | | | |
| | | Cure | | | |
| | | Unresolved | | | |

Remember, any estimation that we do is dependent on historic actuals. The above table shows 2 different estimates:

1. Probability of Outcome category: This is nothing more than avg distribution of population across these categories over the whole time series at the level of months spent in default. Simply put, this value is arrived at by taking all the historic defaults from there 16th month in default snapshot, and their observed resolution category till Max recovery window. For example, if historically 5,000 accounts were incomplete recovery processes with <30% uLTV as of 24th month in default, out of those 5,000 defaults 40% end up being Paid in Full, 20% became REO, 10% were charged off, 25% accounts cured and remaining 5% remained unresolved by the end of 60th month in default. What this tells us is that for an account with <30% uLTV, it has 40% probability of becoming Paid in Full in next 36 months (60-24) and so on.

2. Category level Avg LGDs: This is the historic average LGD of all the Paid in Full accounts, REO accounts and so on. This estimate is at all the historic default level, irrespective of months spent in default. That's why this estimate doesn't

change by months spent in default bucket or uLTV bucket.

So what we have here is a unique estimated LGD value at the level of uLTV and months spent in default. These same estimates can be used when it comes to LGD estimation for Observation Default population.

Table 9: Application of estimated LGDs for Incomplete recovery processes

| Months Spent in Default | uLTV Buckets | Volume | LGD |
|---|---|---|---|
| | <30% | 9 | |
| 22 | 30-60% | 6 | |
| | >60% | 4 | |
| | <30% | 5 | |
| 23 | 30-60% | 3 | |
| | >60% | 3 | |
| | <30% | 4 | 25% |
| 24 | 30-60% | 4 | 34% |
| | >60% | 7 | 42% |

Now we should be able to estimate LGD for 165 perf defaults from Jan 2015 observation and come with up a suitable LGD for the whole month. Another rule of thumb to remember is that any summarization of LGD of defaults/performance defaults from the same month; it should be EAD weighted. While we combine LGDs from different months, it should be default weighted.

With the conclusion of above discussion, one should now be able to give LGD to performance defaults of Observation Non defaults. Next step would be to create the monthly LGD timeseries at the LGD pool level, and then calculate Long run average LGD and Downturn LGD. Max of both these 2 values should be that particular pools LGD parameter to be used for RWA calculation. With this LGD calculation should be doable for all defaults from all vintages.

## 5. Processes specific to Observation Default LGD Calculation

We are not done yet. We are yet to talk about certain specifics related to Observation Defaults.

For this let's revisit our table 1 and now let's break it further. As of Dec 2018, we have 10,000 observation defaults on our books. These accounts have defaulted in past and we are still in the process of recovering money from them. These 10,000 accounts could be at varying stages in Loss mitigation process and may be from different months of default. While there may be some accounts which have recently defaulted in the 2018, there can also be accounts that must have defaulted 5-7 years back and still aren't closed. What should we do? Should all these accounts irrespective of the months spent in default be treated in a similar fashion from an LGD perspective? May be or may be not?

Let's also think about it from LGD modeling/segmentation perspective. Do we think that the kind of variables that help explain LGD of an account with 10 months spent in default are same as the ones relevant to explain LGD of an account with 60

months spent in default? The obvious answer should be highly unlikely, and even if the driver variables are same; their relation and segmentation cut offs may vary. With more months spent in default, there is definitely more information available in our collection's system to help predict LGDs better. For example, you can reasonably expect an account with 10 months spent in default to become Paid in full some time in future, but you may not have the same expectation from an account who has already spent 60 months in default. All these points lead to the introduction of concept of Reference Dates. Simply put, reference dates is a concept of segmenting observation defaults with the months spent in default as first level of segmentation. One may build risk based segmentation on the bottom of this layer. By the way, this is just an introduction and far from over. There are various angles to this concept:

1. How to arrive at appropriate Reference dates

2. Reference Date and LGD Modeling

3. Reference Date and account level LGD estimation for Observation defaults

4. Reference Date and Incomplete recovery processes

**Defining Reference Date:** GL shares 2 ways of doing this:

a) Reference date definition could be event driven. For example, there are several stages in the loss mitigation process of a mortgage account. Collection's team might start with a default customer on a soft note first, like offering him repayment plan/modifications. In early stages, customer might also make some payment promises which he/she might not keep. If none of that works out, we might have to start proceedings on Foreclosure (our legal right to take over collateral in the event of non-payment of dues). We may be at various stages of Foreclosure and then REO (final closure of loan). So, these various check points on the recovery process may become different reference dates. Based on the stage of loss mitigation process a particular observation default is in, might help the bank assign appropriate LGD estimate for that account. So this is one option.

b) Another option is to define reference dates by buckets of months spent in default. 1-12 months could be one bucket; 13-24 could be another and so on. These buckets can be assigned based on LGD differentiation between these buckets with a caveat to have sufficient population in each bucket to do any further statistical segmentation.

There is no preference of choosing one option over the other. It is left to analyst's discretion to pick the one they find suitable for their portfolio.

For more details refer to Paragraph 172, Section 7.1.2 in EBA/GL/2017/16.

Before we go into further details let's assume we choose Time spent in default to arrive at Reference Dates. Since we have Max recovery window length of 60 months, we will have 6 reference date buckets as shown below:

Table 10: Different Reference Date buckets

| Months in Default |
|---|
| 1 – 12 |
| 13 – 24 |
| 25 – 36 |
| 37 – 48 |
| 49 – 60 |
| >60 |

**Reference Date and Modeling:** As already discussed, with change in reference date bucket the relevant predictor variables or their cut offs might change. So, segmentation should be able to capture that aspect. From a data preparation aspect, it should be kept in mind that separate statistical segmentation needs to explored for separate reference date buckets. So, if we were to build a segmentation for 25-36 bucket and Jan15 snapshot is part of the modeling data; we will consider 1,681 observation defaults(refer to table 13) as relevant population (only resolved accounts out of 1,681). We may notice duplication of accounts month over month like the way observed for observation non-defaults. It's recommended to adopt suitable de-duplication strategy to take care of this aspect.

**Reference Date and LGD estimation:** This piece is interesting. Let's begin with an example. Let's assume an account X defaulted in Oct 2016 with a balance of $80,000. Our collection's team was able to convenience customer to make some payments, but the customer still has not cured. As of Dec 2018, this account is part of observation default population with an outstanding balance of $72,000. Now, our job is to set aside capital for this account. Which balance should we use as EAD for this account? $80,000 or $72,000? Might seem like a stupid question. Off course, $72,000 since that is the value at risk standing Dec18. $8,000 has already been recovered. Now, let's think about this? With the reception of payment, should the associated LGD for this account also change? Not sure.

Let's look at this example below with continuous calculation of LGD. This account Y defaulted in Jan2015 with a balance of $10,000. It takes 30 months to resolve with a total recovery of $8,000 over a number of smaller payments. For the sake of simplicity, let's keep aside cost and discounting for now. The full performance LGD for this account is 20%. Look at the table carefully.

Table: 11 Monthly updating EAD and continuous calculation of LGD

| Month | Months in Default | EAD | New EAD | Recovery | Cumulative Recovery | New outstanding recovery | LGD |
|---|---|---|---|---|---|---|---|
| Jan-15 | 1 | $ 10,000 | $ 10,000 | $ 0 | $ 0 | $ 8,000 | 20% |
| Feb-15 | 2 | $ 10,000 | $ 10,000 | $ 0 | $ 0 | $ 8,000 | 20% |
| Mar-15 | 3 | $ 10,000 | $ 9,800 | $ 200 | $ 200 | $ 7,800 | 20% |
| Apr-15 | 4 | $ 10,000 | $ 9,600 | $ 200 | $ 400 | $ 7,600 | 21% |

| Month | Months in Default | EAD | New EAD | Recovery | Cumulative Recovery | New outstanding recovery | LGD |
|---|---|---|---|---|---|---|---|
| May-15 | 5 | $ 10,000 | $ 9,400 | $ 200 | $ 600 | $ 7,400 | 21% |
| Jun-15 | 6 | $ 10,000 | $ 9,200 | $ 200 | $ 800 | $ 7,200 | 22% |
| Jul-15 | 7 | $ 10,000 | $ 9,200 | $ 0 | $ 800 | $ 7,200 | 22% |
| Aug-15 | 8 | $ 10,000 | $ 9,200 | $ 0 | $ 800 | $ 7,200 | 22% |
| Sep-15 | 9 | $ 10,000 | $ 8,700 | $ 500 | $ 1,300 | $ 6,700 | 23% |
| Oct-15 | 10 | $ 10,000 | $ 8,700 | $ 0 | $ 1,300 | $ 6,700 | 23% |
| Nov-15 | 11 | $ 10,000 | $ 8,700 | $ 0 | $ 1,300 | $ 6,700 | 23% |
| Dec-15 | 12 | $ 10,000 | $ 8,700 | $ 0 | $ 1,300 | $ 6,700 | 23% |
| Jan-16 | 13 | $ 10,000 | $ 7,700 | $ 1,000 | $ 2,300 | $ 5,700 | 26% |
| Feb-16 | 14 | $ 10,000 | $ 7,700 | $ 0 | $ 2,300 | $ 5,700 | 26% |
| Mar-16 | 15 | $ 10,000 | $ 7,700 | $ 0 | $ 2,300 | $ 5,700 | 26% |
| Apr-16 | 16 | $ 10,000 | $ 7,700 | $ 0 | $ 2,300 | $ 5,700 | 26% |
| May-16 | 17 | $ 10,000 | $ 6,200 | $ 1,500 | $ 3,800 | $ 4,200 | 32% |
| Jun-16 | 18 | $ 10,000 | $ 6,200 | $ 0 | $ 3,800 | $ 4,200 | 32% |
| Jul-16 | 19 | $ 10,000 | $ 6,200 | $ 0 | $ 3,800 | $ 4,200 | 32% |
| Aug-16 | 20 | $ 10,000 | $ 6,200 | $ 0 | $ 3,800 | $ 4,200 | 32% |
| Sep-16 | 21 | $ 10,000 | $ 6,200 | $ 0 | $ 3,800 | $ 4,200 | 32% |
| Oct-16 | 22 | $ 10,000 | $ 6,000 | $ 200 | $ 4,000 | $ 4,000 | 33% |
| Nov-16 | 23 | $ 10,000 | $ 6,000 | $ 0 | $ 4,000 | $ 4,000 | 33% |
| Dec-16 | 24 | $ 10,000 | $ 6,000 | $ 0 | $ 4,000 | $ 4,000 | 33% |
| Jan-17 | 25 | $ 10,000 | $ 6,000 | $ 0 | $ 4,000 | $ 4,000 | 33% |
| Feb-17 | 26 | $ 10,000 | $ 6,000 | $ 0 | $ 4,000 | $ 4,000 | 33% |
| Mar-17 | 27 | $ 10,000 | $ 6,000 | $ 0 | $ 4,000 | $ 4,000 | 33% |
| Apr-17 | 28 | $ 10,000 | $ 6,000 | $ 0 | $ 4,000 | $ 4,000 | 33% |
| May-17 | 29 | $ 10,000 | $ 6,000 | $ 0 | $ 4,000 | $ 4,000 | 33% |
| Jun-17 | 30 | $ 10,000 | $ 2,000 | $ 4,000 | $ 8,000 | $ 0 | 100% |

As of Jan 2015 observation, this account had $10,000 outstanding and we were able to recover $8,000 from this account in 30 months of performance. So, standing Jan 2015 and looking forward, this account gave us a loss of 20% of its EAD as of Jan 2015.

Same is true for the month Feb 2015. As of month Sep 2015, this same account has made us payments worth $1,300. So, as of Sep 2015 our outstanding balance or EAD is $8,700 and we loose $2,000 out of this $8,700. So technically speaking, now our forward looking LGD for this account is 23% of our outstanding EAD of Sep 2015. I hope you are able to connect it with the real time implementation environment. An LGD estimate is the loss as percentage of EAD. With the concept of continuous update of EAD in real time environment, ideally the Loss expectations should also vary.

As of month May 2016, this account owed us $6,200 and we lost $2,000 out of that

$6,200. So our realistic LGD estimate as of May 2016 is 32% of our EAD. This happens because most of the losses are booked at last.

EAD for an account is auto updated on receipt of any recovery, but changing LGD every month based on recoveries might be a cumbersome task. This is where reference date comes to rescue. In the above example, we updated EAD and connected recovery on a monthly basis. What if, instead of doing it every month, we update it as and when reference date bucket changes. So, earlier when we had assumed 1-12, 13-24, etc. as reference dates, we will update EAD and recovery values as of 1st, 13th, 25th month and so on. What this means is with the change in reference date, we update our recoveries, costs and EADs to the one realised from start of that window. For our account level LGD calculation, we consider EAD as of 13th month and all the recoveries and costs that were incurred after that month. This gives us a more realistic estimate of LGD that can be tied back to implementation scenario as well. Table below illustrates our view point.

Table 12: Account level LGD using Reference Date buckets

| Reference Date Bucket | EAD (as of 1st month of that reference window) | Forward looking recovery | LGD |
|---|---|---|---|
| 1-12 | $ 10,000 | $ 8,000 | 20% |
| 13-24 | $ 7,700 | $ 5,700 | 26% |
| 25-36 | $ 6,000 | $ 4,000 | 33% |

Along with this change, there is one more perspective of discounting. In case of observation non-defaults, all the costs and recoveries were discounted back to the moment of default. In case of LGD estimation for observation defaults, with the update of reference date; the snapshot of discounting also gets updated. For example, for the LGD calculation of an account as of 13th month in default will now have all the costs and recoveries post 13th month discounted back to 13th month and not to the month of default. Refer to Paragraph 179, Section 7.3.1 of EBA/GL/2017/16.

Remember, all the LGD parameter estimation is based on actual historic LGDs. This account Y that we just talked about is a historic resolved (closed recovery process) account. So, this account will form part of observation default population in LGD timeseries from Jan15 to Jun17.

Table 13: Use of account level LGD in monthly LGD calculation after incorporating Reference date concept

| Month | # of Observation defaults | Months Spent in Default | Account Y present | Contributed LGD of Account Y |
|---|---|---|---|---|
| Nov-14 | 7,786 | - | 0 | - |
| Dec-14 | 7,122 | - | 0 | - |
| Jan-15 | 7,166 | 1 | 1 | 20% |
| Feb-15 | 7,940 | 2 | 1 | 20% |
| Mar-15 | 7,838 | 3 | 1 | 20% |
| Apr-15 | 7,928 | 4 | 1 | 20% |

| Month | # of Observation defaults | Months Spent in Default | Account Y present | Contributed LGD of Account Y |
|---|---|---|---|---|
| May-15 | 7,199 | 5 | 1 | 20% |
| Jun-15 | 7,262 | 6 | 1 | 20% |
| Jul-15 | 7,724 | 7 | 1 | 20% |
| Aug-15 | 7,078 | 8 | 1 | 20% |
| Sep-15 | 7,937 | 9 | 1 | 20% |
| Oct-15 | 7,324 | 10 | 1 | 20% |
| Nov-15 | 7,500 | 11 | 1 | 20% |
| Dec-15 | 7,195 | 12 | 1 | 20% |
| Jan-16 | 7,530 | 13 | 1 | 26% |
| Feb-16 | 7,800 | 14 | 1 | 26% |
| Mar-16 | 7,166 | 15 | 1 | 26% |
| Apr-16 | 7,815 | 16 | 1 | 26% |
| May-16 | 7,894 | 17 | 1 | 26% |
| Jun-16 | 7,447 | 18 | 1 | 26% |
| Jul-16 | 7,552 | 19 | 1 | 26% |
| Aug-16 | 7,337 | 20 | 1 | 26% |
| Sep-16 | 7,340 | 21 | 1 | 26% |
| Oct-16 | 7,068 | 22 | 1 | 26% |
| Nov-16 | 7,806 | 23 | 1 | 26% |
| Dec-16 | 7,185 | 24 | 1 | 26% |
| Jan-17 | 7,947 | 25 | 1 | 33% |
| Feb-17 | 7,934 | 26 | 1 | 33% |
| Mar-17 | 7,530 | 27 | 1 | 33% |
| Apr-17 | 7,885 | 28 | 1 | 33% |
| May-17 | 7,362 | 29 | 1 | 33% |
| Jun-17 | 7,036 | 30 | 1 | 33% |
| Jul-17 | 7,488 | - | 0 | - |
| Aug-17 | 7,873 | - | 0 | - |
| Sep-17 | 7,492 | - | 0 | - |

So, in the table above, it is required that all the 7,262 defaulted accounts as of Jun 2015 need to have an actual/estimated LGD value to be able to arrive at Jun 2015's overall monthly LGD estimate. Out of these 7,262 defaults, accounts could be from varying months spent in default. Similarly, some of them might have resolved in the performance available until Dec 2017. Some of them, might be unresolved and some others left over would be incomplete recovery processes. For all the accounts resolved in performance, we will use their actual LGDs. For the ones, that were able to get more than 60 months of performance and still remained open are classified as Unresolved accounts and receive their LGDs from unresolved framework. The accounts that remain open as of Dec2017 and haven't got full 60 months of performance are classified as incomplete recovery processes and receive their LGDs from

Incomplete recovery framework. Our account Y is part of resolved population as of Jun 2015. The question is what should be the representative LGD booked as a contribution of this account for Jun 2015?

As of Jun 2015, account has spent 6 months in default and thus is classified as 1-12 reference date account. For LGD calculation as of Jun 2015, we take EAD as of Jan 2015, and consider all the costs incurred and recoveries received post Jan 2015 till Dec 2017. All these cash flows are discounted to the month of default Jan 2015.

As of Jun 2016, this account has spent 18 months in default. The new reference date for this account is the start month of reference window (13-24) i.e. Jan 2016. So, in order to calculate representative LGD we will use EAD value as of Jan 2016 ($7,700) and all the subsequent costs and recoveries are discounted back to Jan 2016 and not 2015.

In this way, in case of observation defaults an account's representative LGD for LGD estimation might change with the changing month. It might seem absurd to have an account representing different LGD numbers for itself at different snapshots in time, but this view is in line with the changing EAD value from an implementation perspective. Refer to Pg 35-36 (Section Reference Dates) in EBA/GL/2017/16.

Lastly let's talk about impact of **reference date on incomplete recovery processes:**

GL recommends that accounts which are incomplete recovery processes should be considered as part of LGD timeseries only if they have actual performance available until the end of next reference date. Non-compliance to this guideline will lead to circular referencing of same account in estimation and implementation. This may sound very complex and undecipherable. Let's understand this with an example. Let's assume there is an account 'Z' which defaults in the month of May2015. So as of Jun 2015, it has spent one month in default and so on. This account is an open defaulted account as of Dec 2017 (end of performance and month of estimation). Since as of Dec2017, this account has got only 31 months' worth of performance in default, which is less than 60 months of max recovery window; this account will be classified as incomplete recovery process and will receive its representative LGD from incomplete recovery framework. With this information in mind, let's pay attention to below illustration now.

Table 14: Movement of an account across different reference date buckets during LGD Quantification

| Month | Reference Date Buckets | | | | | | |
|---|---|---|---|---|---|---|---|
| | 1-12 | 13-24 | 25-36 | 37-48 | 49-60 | >60 | Total |
| Jan-15 | 2,030 | 1,818 | 1,681 | 1,438 | 1,176 | 1,104 | 9,247 |
| Feb-15 | 2,178 | 2,182 | 1,519 | 1,684 | 1,094 | 794 | 9,451 |
| Mar-15 | 2,445 | 2,037 | 1,647 | 1,393 | 1,391 | 1,246 | 10,159 |
| Apr-15 | 2,057 | 1,926 | 1,571 | 1,639 | 1,377 | 844 | 9,414 |
| May-15 | 2,161 | 2,226 | 1,703 | 1,706 | 1,034 | 1,072 | 9,902 |
| Jun-15 | 2,028 | 2,225 | 1,995 | 1,703 | 1,161 | 1,017 | 10,129 |
| Jul-15 | 2,021 | 2,139 | 1,577 | 1,456 | 1,144 | 1,238 | 9,575 |
| Aug-15 | 2,270 | 1,836 | 1,587 | 1,604 | 1,043 | 1,201 | 9,541 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Sep-15 | 2,331 | 2,037 | 1,972 | 1,617 | 1,226 | 903 | 10,086 |
| Oct-15 | 2,417 | 1,821 | 1,506 | 1,358 | 1,330 | 917 | 9,349 |
| Nov-15 | 2,252 | 1,803 | 1,689 | 1,327 | 1,489 | 956 | 9,516 |
| Dec-15 | 2,162 | 2,172 | 1,571 | 1,601 | 1,463 | 821 | 9,790 |
| Jan-16 | 2,426 | 2,118 | 2,243 | 1,900 | 1,631 | 1,093 | 11,411 |
| Feb-16 | 2,411 | 2,110 | 1,801 | 1,541 | 1,704 | 1,192 | 10,759 |
| Mar-16 | 2,332 | 2,025 | 2,213 | 1,749 | 1,263 | 1,465 | 11,047 |
| Apr-16 | 2,716 | 2,107 | 1,870 | 1,595 | 1,518 | 1,253 | 11,059 |
| May-16 | 2,440 | 2,499 | 2,047 | 1,910 | 1,705 | 1,015 | 11,616 |
| Jun-16 | 2,286 | 2,452 | 2,144 | 1,941 | 1,575 | 1,383 | 11,781 |
| Jul-16 | 2,482 | 2,068 | 1,917 | 1,820 | 1,516 | 1,342 | 11,145 |
| Aug-16 | 2,416 | 2,351 | 2,147 | 1,577 | 1,469 | 1,484 | 11,444 |
| Sep-16 | 2,335 | 2,450 | 2,130 | 1,551 | 1,342 | 1,251 | 11,059 |
| Oct-16 | 2,353 | 2,088 | 2,016 | 1,774 | 1,266 | 1,303 | 10,800 |
| Nov-16 | 2,449 | 2,138 | 2,052 | 1,638 | 1,541 | 1,248 | 11,066 |
| Dec-16 | 2,560 | 2,202 | 2,127 | 1,722 | 1,254 | 1,067 | 10,932 |
| Jan-17 | 2,939 | 2,322 | 2,208 | 2,154 | 1,505 | 1,253 | 12,381 |
| Feb-17 | 2,792 | 2,325 | 2,336 | 1,816 | 1,888 | 1,418 | 12,575 |
| Mar-17 | 2,656 | 2,611 | 2,030 | 1,834 | 1,879 | 1,695 | 12,705 |
| Apr-17 | 2,757 | 2,351 | 2,062 | 2,172 | 1,826 | 1,369 | 12,537 |
| May-17 | 2,510 | 2,276 | 2,232 | 2,072 | 1,571 | 1,426 | 12,087 |
| Jun-17 | 2,984 | 2,534 | 2,030 | 1,789 | 1,685 | 1,427 | 12,449 |
| Jul-17 | 2,868 | 2,323 | 2,156 | 2,128 | 1,500 | 1,526 | 12,501 |
| Aug-17 | 2,989 | 2,516 | 2,391 | 1,817 | 1,829 | 1,364 | 12,906 |
| Sep-17 | 2,676 | 2,351 | 2,349 | 1,858 | 1,528 | 1,611 | 12,373 |
| Oct-17 | 2,619 | 2,739 | 2,454 | 1,953 | 1,670 | 1,613 | 13,048 |
| Nov-17 | 2,848 | 2,541 | 2,257 | 1,806 | 1,942 | 1,718 | 13,112 |
| Dec-17 | 2,992 | 2,383 | 2,454 | 1,944 | 1,889 | 1,343 | 13,005 |

The above table is a depiction of observation default timeseries. This timeseries assumes that segmentation was performed only at a level of reference date. In reality, there might a subsequent layer of statistical segmentation driven by attributes like uLTV (or some other variable), but that view has been avoided for the sake of simplicity. The illustration shows distribution of observation defaults across various months in default buckets. For example, as of Jul 2015 there were in total 9,575 defaulted accounts on our books. These accounts were distributed into various months spent in default buckets as shown in the illustration. As seen there were 1577 defaults which were on our books as of Jul 2015 and had already spent 25-36 months in default. Out of these 1577 accounts, 906 of them resolve in the performance available till Dec 2017. 427 of them are incomplete recovery processes and remaining 244 accounts are unresolved. As explained on multiple occasions, these 906 accounts will contribute their actual LGDs in LGD timeseries, 427 will leverage their LGDs from Incomplete recovery framework and 244 from Unresolved framework. Our account 'Z' is sitting as one of the 2021 accounts part of 1-12 bucket. It is an in-

complete recovery process account. In the illustration, you can notice how this account migrates to different LGD bucket in Jun 2016 because of the change in months spent in default from 12 to 13.

As per the guidelines, it is fine to include this account 'Z' as part of LGD quantification time series until May 2017 snapshot but not after that. The reason behind this objection is that if we were to calculate RWA using this data for the month of Jan 2018, this account 'Z' will sit in 25-36 bucket as of implementation, since it's a default. Based on this status, it will use LGD parameter for 25-36 bucket calculated by taking Long run average of LGDs of accounts part of timeseries from Feb 2007 to Dec 2017. In our estimation process, if we include this account in 25-26 bucket for the month of Jun 2017 till Dec 2017, then this account 'Z' would have participated in the LGD estimation process as well and would leverage the same LGDs for assigning it capital in the month of Jan 2018. This has led to circular reference of this account, where LGD estimates of the same account were used to derive its forward looking LGD. This is considered as irrational. Hence, an incomplete recovery process is included in timeseries until the point it has full performance available till the end of that reference window; else not.

## 6. Conclusion

With all the information shared above, combined with the information from discussion regarding Observation Non-Defaults, one should be able to calculate account level LGD for all observation defaults and non-defaults as well and subsequently create the LGD timeseries for LGD estimation. Principles related to Unresolved LGD and Incomplete recovery processes stay the same for Observation default and non-default view. Where there are any changes, it has already been covered in the specifics for observation default section.

Just to summarize, our objective is to calculate LGD parameter at pool level. First step is to calculate account level LGDs for all the resolved accounts, whether inside or outside max recovery window. We should be able to calculate their actual LGDs. These LGDs might change with changing snapshot for Observation defaults because of shift in reference window. For defaults which have not resolved, these could either be unresolved or incomplete recovery processes. We use respective framework to calculate estimated LGDs for these accounts. Now we have LGD value for all the historic defaults. We can use only resolved accounts to build model/segmentation separately for Observation defaults and non-defaults. For Observation defaults we need to take care of reference dates as well. Finally, identify downturn period, calculate Long-run average LGD & Downturn LGD. Take max as the LGD parameter for capital calculation.

## 7. Authors and Affiliations

Naval Chaudhary is a risk management professional at a financial institution. He has more than six years of experience in credit risk model development, regulatory capital estimation and regulatory framework design.

Manish Malik is a credit risk professional with 12 years of experience building credit risk AIRB and scoring models.

## References

1. EBA/GL/2017/16 – Guidelines on PD estimation, LGD estimation and the treatment of defaulted exposures
2. EBA/GL/2016/07 – Guidelines on the application of the definition of default under Article 178 of Regulation (EU) No 575/2013
3. SS 11/13 – Supervisory Statement Internal Rating based approaches